VOL. XX JUNE, 1936 No. 3

THE JOURNAL OF APPLIED PSYCHOLOGY

JAMES P. PORTER, *Editor*
OHIO UNIVERSITY ATHENS, OHIO

With a Board of Cooperating Editors
and
PAUL S. ACHILLES HENRY C. LINK
Associate Editors Representing
THE PSYCHOLOGICAL CORPORATION

SOME INTERESTING CONTRIBUTIONS IN THIS ISSUE:

EDWARD L. THORNDIKE. The Value of Reported Likes and Dislikes for Various Experiences and Activities as Indications of Personal Traits

HAROLD V. GASKILL and RICHARD L. HOLCOMB. The Effectiveness of Appeal in Radio Advertising: A Technique with Some Typical Results

ROSS STAGNER and E. T. KATZOFF. Personality as Related to Birth Order and Family Size

BRUCE D. GREENSHIELDS. Reaction Time in Automobile Driving

E. G. WILLIAMSON. The Rôle of Faculty Counseling in Scholastic Motivation

PUBLISHED BI-MONTHLY BY THE EDITOR
OHIO UNIVERSITY, ATHENS, OHIO

PRINTED BY THE SCIENCE PRESS PRINTING COMPANY, LANCASTER, PA.

Entered as second-class matter at the post office at Athens, Ohio, under the act of March 3, 1879. Additional entry May 7, 1935, at the post office at Lancaster, Pa. Copyright 1936 by James P. Porter.

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. F. Book, Indiana University, W. V. Bingham, Personnel Research Federation, New York; M. E. Broom, State Teachers College, San Diego, Calif.; H. E. Burtt, Ohio State University; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; J. E. W. Wallin, Atlantic University; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry M. McComas, Johns Hopkins University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, New York; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; Joseph Jastrow, New School of Social Research, New York; G. M. Whipple, Danvers, Mass.; Helen Thompson Woolley, Teachers College, Columbia University.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and offprints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this Journal; in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published bi-monthly. The subscription price is $6.00 a year. Single numbers $1.25 to $2.00, depending upon the number of pages. Remittances and business communications should be addressed to the Editor, James P. Porter, Ohio University, Athens, Ohio.

Vol. XX JUNE, 1936 No. 3

THE JOURNAL OF APPLIED PSYCHOLOGY

Edited by
JAMES P. PORTER, Ohio University
Athens, Ohio

TABLE OF CONTENTS

EDWARD L. THORNDIKE. The Value of Reported Likes and Dislikes for Various Experiences and Activities as Indications of Personal Traits 285

E. G. WILLIAMSON. The Rôle of Faculty Counseling in Scholastic Motivation 314

HAROLD V. GASKILL AND RICHARD L. HOLCOMB. The Effectiveness of Appeal in Radio Advertising: A Technique with Some Typical Results 325

ROSS STAGNER AND E. T. KATZOFF. Personality as Related to Birth Order and Family Size 340

JAMES H. TAYLOR. The Relationship Between Finger Length, Hand Width and Musical Ability 347

BRUCE D. GREENSHIELDS. Reaction Time in Automobile Driving 353

RALPH M. STOGDILL. The Measurement of Attitudes Toward Parental Control and the Social Adjustments of Children 359

MARY M. ROOS. Sampling Theory as Used in the Determination of Psychological Trends in Volume Mail 368

IRVING LORGE. A Table of Percentile Equivalents for Eight Intelligence Tests Frequently Used with Adults 392

HARRY PELLE HARTKEMEIER AND LAURA MAE BROWN. Multiple Correlation and the Multiple Factor Method 396

HOWARD P. LONGSTAFF. Preliminary Results of a Study of Mothers' Opinions of Children's Radio Programs 416

NEWS AND NOTES 420

BOOK REVIEWS 422

NEW BOOKS AND PAMPHLETS 426

ANNOUNCEMENT

Arrangements effective with this issue have been made whereby The Psychological Corporation will participate in the administration of this Journal through membership on the editorial board of Dr. Paul S. Achilles and Dr. Henry C. Link as associate editors representing the Corporation. It is hoped that these arrangements will further the advancement of scientific knowledge of the useful applications of psychology.

THE VALUE OF REPORTED LIKES AND DISLIKES FOR VARIOUS EXPERIENCES AND ACTIVITIES AS INDICATIONS OF PERSONAL TRAITS

EDWARD L. THORNDIKE
Teachers College, Columbia University

IF a person reports honestly his degree of liking or disliking for each of a list of activities and experiences such as those of the Strong vocational interest blank or of other concrete lists, or of the list parts of which are printed below, his report should give indications of his status in fundamental qualities of personal make-up, if there are such, and if we use his report wisely.

The lists which will be studied here comprised 250 items, 187 items, and 150 items, concerning each of which certain persons reported by symbols (1, 2, 5, 8, and 9) meaning

(1) I am sure that I like it
(2) I think I like it but am not sure
(5) I neither like nor dislike it
(8) I think I dislike it, but am not sure
(9) I am sure that I dislike it

We also had by the courtesy of Dr. Lorge three complete records for each person with the Strong Vocational Interest Blank.

The items included some rather broad categories and some very specific activities and experiences.

One way to use such reports would be to have an adequate sample of persons make reports, compute the intercorrelations of the items, and the correlation of each item with various criteria in the shape of other known facts about the persons, and perhaps apply a technique of Kelley, Thurstone or Hotelling to the factorial analysis of all these. Practical considerations veto this task for a thousand or even a hundred

items. So we begin by grouping items. In proportion as the items that we put together as presumably (that is to say, in our judgment) related by a fair degree of common causation continue to show substantial individual differentiation after summing, and substantial self correlation between two random halves of each sum, our grouping is justified. We may then proceed to check the summed scores for groups of items with such criteria as are available.

In groups 7 to 35 (Nos. 1 to 6 refer to measures to be described later) we never used an item in more than one group, because we wished, in case the records gave sufficiently reliable and valid measures of interests, to have each of these measures derived from independent sources. We also left many items unassigned to any group, in the hope of later checking them against groups of known significance before placing them.

Our groupings were as shown below for some of the items in the lists of 250 and 187. The items in the list of 150 and in the Strong blank were used (1) as an independent check, (2) to determine reliability coefficients, and (3) for certain special purposes and need not be shown here.

In Groups 36 to 42 there was some duplication of items *inter se* and with item-groups 7 to 35.[1] In 36 (called

[1] The Reality Score (36) was due to 122 items, of which 9 were used also for score 7, 2 for score 8, 4 for score 9, 1 for 12, 1 for 13, 2 for 15, 11 for 16, 4 for 17, 2 for 18, 4 for 19, 1 for 21, 3 for 22, 2 for 24, 4 for 25, 1 for 26, 1 for 27, 9 for 29, 3 for 31, 3 for 32, 9 for 34, and 1 for 35.

The Imagination Fiction Score (38) was due to 48 items, of which 6 were used also for score 11, 2 for score 15, 3 for score 21, and 1 each for scores 7, 10, 13, 14, 16, 17, 18, 23, 25, 28, 32, and 35.

The Talk Score (37) was due to 27 items, of which 6 were used also in score 8, 1 in score 9, 1 in 21, 1 in 22, 3 in 23, 3 in 24, 3 in 29, 1 in 31, and 4 in 32.

The Indulgence Score (39) was due to 21 items, of which 3 were used in 29 and one each in 28, 32 and 42.

The Moral Score (40) was due to 32 items, of which 8 were used also for 23, 2 for 31, and 1 each for 16, 17, 21 and 22.

'Reality') were included items of activity or experience concerned exclusively or chiefly with real things and people. In 38 (called Imagination or Fiction) were items such as (from the 187 set): Spend an afternoon in an art gallery (also in 13), Spend an afternoon in a library (also in 11), Make a card catalog (also in 17), Copy quotations and references (also in 17), Make puns (also in 16). In 37 (called Talk) were items where speech by the person was prominent. In 39 (called Indulgence) and 40 (called Moral) were respectively items commonly thought of as done chiefly for pleasure and done chiefly for duty. Group 41 included only Sedentary Games and 42 included only Dancing.

The amount of duplication is reported in Appendix I. Whenever it may have influenced any conclusion the fact will be stated.

ITEM–GROUPS

In the 250 set, items numbered from 21–50 refer to school life, at age 13–19; items 51–100 refer to life outside school at age 13–19; items 101–250 refer to age 20–25. All items in the 187 set refer to age 20–25.

7 Things.

250 set.
38. Manual training
39. Shop-work
40. Trade work of any sort
46. Cooking
47. Domestic science (in general)
119. Repairing an automobile
124. Working with tools
187. Sharpen a knife
188. Take a clock apart and put it together
225. Do picture puzzles
227. Do mechanical puzzles

187 set.
4. Be a cabinet maker
9. Be an electrician
10. Be an inventor
12. Be a plumber
14. Be a repairer of typewriters
37. Spend an afternoon visiting a shoe factory
52. Make ship models

The Sedentary Games Score (41) was due to 3 items, of which 2 were used in 33.

The Dancing Score (42) was due to 5 items, of which one was used in 29.

106. Clean a bicycle
113. Oil a sewing machine

8 People.

250 set.
104. Bargaining (also in item-group 22)
105. Selling to a customer
127. Going on errands
186. Go around applying for a job
241. Working as a guide on a sightseeing trip

187 set.
1. Be a book agent
2. Be a life insurance agent
3. Be an ambassador or diplomat
6. Be a chief clerk in an office
7. Be a conductor on a railroad
13. Be a politician
16. Be a secretary to a senator
17. Be a social secretary to a society lady
26. Read the personal column in a newspaper
31. Spend an afternoon making social calls
58. Take pictures of insane people

9 Animals.

250 set.
33. Studying zoology
56. Taking care of animals
174. Play with a dog or cat
192. Feed the horses and cows
193. Feed the hens and chickens
200. See a beautiful horse

187 set.
43. Teach a parrot to talk
44. Teach a dog tricks
56. Take pictures of domestic animals
57. Take pictures of wild animals
68. Have a pet canary
69. Have a pet monkey
86. Washing a dog
116. Raise fancy pigeons

10 Words.

250 set.
23. Latin
24. French, German or Italian
217. Try to write a poem
226. Do cross-word puzzles

187 set.
25. Read a page of a dictionary
51. Make puns
53. Make slang words or novel expressions

11 Reading.

250 set.
25. English literature
57. Reading stories (in elem. school)
128. Reading stories (20 to 25)
130. Reading serious books

187 set.
29. Spend an afternoon in a library
83. Reading in bed

84. Reading on a train
85. Reading while music is being played
141. Learn that if you can study for three years you can have a great career as a writer

12 Bodily Exercise.

250 set.
43. Gymnastics
51. Athletics (in teens)
64. Doing heavy work, chopping, digging, etc.
74. Swimming
138. Athletic sports not competitive: hiking, bicycling, hunting, skating
169. Take a walk in the city

187 set.
82. Wrestling
126. Walking to and from work or school
146. Do a daily dozen of physical exercises
160. Saw wood

13 Art.

250 set.
41. Art (in elem. school)
67. Drawing (in teens)
68. Painting (in teens)
132. Drawing (20 to 25)
133. Painting (20 to 25)

187 set.
28. Spend an afternoon in an art gallery
38. Spend an afternoon watching a famous painter paint
39. Spend an afternoon watching a famous sculptor model
59. Take pictures of landscapes
129. Spend $500 for a friend for pictures for his home
136. See an exhibit of modern French paintings
183. Try to learn to make etchings.

14 Music.

250 set.
42. Music
69. Singing
70. Playing piano, violin or other instrument
88. Going to concerts
134. Singing (20 to 25)
136. Listening to music
179. Go to a good concert

187 set.
30. Spend an afternoon listening to Chinese music with explanations
63. Go to an organ recital
64. Go to a concert by a college glee club
79. Listening to bird songs
135. Take a trip to Mexico to collect phonographic records of Indian music
140. Learn that if you can study for three years you can have a great career as a singer
182. Try to learn to play the flute

15 Beauty.

250 set.
58. Reading poetry (in teens)
63. Doing fancy sewing or embroidery (in teens)
114. Doing fancy sewing or embroidery (20 to 25)
129. Reading poetry (20 to 25)
197. Seeing a beautiful sunset
198. Seeing a beautiful painting
199. Seeing a beautiful dancer
236. Visit the art museum

187 set.
8. Be a dress designer
22. Read books on lace and embroidery
40. Watch an expert appraise antiques
47. Make a collection of 50-cent objects that are really beautiful
130. Spend $20,000 in making a collection of rugs for a museum
131. Spend $40,000 in making a collection of pottery for a museum
134. Take a trip through Eastern Europe buying old furniture, ikons, etc., for a museum
137. See an exhibit of Chinese porcelain
147. Earn your living as a jewel expert

16 Responsibility.

250 set.
110. Showing others how to do their work
161. Cook a fancy dinner
221. Plan for a party
245. Working on a committee to secure money for a good cause

187 set.
19. Be a vocational counselor
88. Be a buyer (i.e. head of department) of china and glass for a big department store
89. Be a buyer (i.e. head of department) of dress goods for a big department store
90. Be a buyer (i.e. head of department) of furniture for a big department store
91. Be a buyer (i.e. head of department) of sporting goods for a big department store
93. Be mayor of a small city
98. Manage a small bookstore of your own
99. Manage a small drugstore of your own
100. Manage a small grocery store of your own
101. Manage a small factory of your own
102. Manage a small farm of your own
117. Read manuscripts and advise a publisher whether to publish them

17 Detail.

250 set.
66. Washing dishes (in teens)
125. Washing dishes
126. Making beds
191. Darn stockings

187 set. 74. Addressing letters
108. Copy quotations and references
161. Sort mail

18 System.
250 set. 182. Clean and arrange your bureau
183. Put everything in order in a room or office

187 set. 46. Make a card catalog of persons, books, etc.

19 Neatness.
250 set. 62. Doing plain sewing (in teens)
106. Filing papers
123. Cleaning house
189. Mend clothes
190. Shine your shoes
223. Clean up after a party (given by you)

187 set. 18. Be a street-cleaner
72. Have your hands manicured
75. Brushing your clothes
76. Brushing your hair
111. Keep neat and accurate accounts
119. Set a table
144. Clean up a yard or garden
155. Polish brass on a boat or car

20 Sum of scores in 17, 18, and 19.

21 Approval.
250 set. 201. Rescue a child from danger, and have the crowd cheer you
202. Stop a panic in a theater that gets on fire, and be praised in the newspapers
206. (Reversed). Have a waiter look sour because you did not tip him
208. (Reversed). Have your boss blame you when you didn't deserve it
212. Be elected to an office
213. Win a prize at a party
214. Be praised by your boss
219. Tell a funny story
229. Think what you would do if you were famous

187 set. 77. (Reversed). Having a crowd of beggars jeer at you because you give them nothing
109. Have clerks in stores comment on your good taste and judgment
118. Receive applause when you speak at a meeting
148. Find that your opinions on a subject are held also by noted writers
149. Get mentioned in the Atlantic Monthly as a promising poet
150. Have children admire your skill in swimming, diving, gymnastics, etc.

151. Have men in a garage admire your car
152. Help a rich drunkard throw money for a crowd to catch
169. Be surrounded by a crowd who mistake you for a popular actor or actress
174. Have people compliment you on your new clothes
179. Tackle a thief who is running away and hold him till a policeman comes

22 Mastery or domination over persons.

250 set.
104. Bargaining
111. Keeping others at work
204. (Reversed). Have people give you orders who have no right to
205. (Reversed). Have people give you advice when you did not ask for it
209. Make a cruel man stop beating a horse
210. Make a bully stop teasing a child
211. Call down somebody who is fresh to you

187 set.
48. Make a person who has insulted one of your friends apologize to her
96. Be strong enough to knock down anybody who pushes you on the street, and do so
104. Boss a gang of boys and make them do useful work
105. Catch a man stealing from a child and make him give back the thing plus a dollar
154. (Reversed). Let a policeman bawl you out without answering back
162. Stop a landlord from evicting a poor family with a sick father, by bluffing him into the fear than you and your gang will beat him up
175. Lead a disturbance against a militaristic lecturer and break up the meeting
176. (Reversed). Obey a teacher to save making a fuss although the teacher is in the wrong
178. Stop boys who are fighting in the street and make them go home

23 Religion.

250 set.
84. Going to church
155. Going to church (20 to 25)
172. Attend a religious service.
173. Read a religious book (Bible, Talmud, etc.)

187 set.
5. Be a clergyman
11. Be a missionary
15. Be a secretary to a bishop
45. Teach boys of sixteen in Sunday School
115. Pray for the sick and afflicted
120. Sing hymns
158. Read the Bible or Talmud to a poor blind man

24 Display; Public Attention.

250 set.
45. Dramatics (in high school)

102. Getting dressed for a party
203. Have people stare at you because your clothes are queer
207. Have people gossip about you
218. Make a speech to an audience of 100 people
220. Take part in a play
242. Be usher or bridesmaid at a wedding
246. Have your teacher tell the other pupils that you are a model pupil
247. Have your boss tell the other employees that you are best

187 set.
73. Have your picture taken
78. Introducing a speaker at a meeting
92. Be a head-waiter
97. Be toastmaster at a banquet
103. Advertise a new style of bathing suit by appearing in a store window wearing it.
112. Lead a procession
143. Address a meeting
157. Read aloud to a group
159. Rise up in a meeting and say "That is a lie" to a person who has uttered an outrageous lie

25 Novelty.
250 set.
85. Going on automobile trips
131. Reading the newspaper
139. Traveling and seeing new places
163. Eat dog biscuit
164. Eat a live beetle
178. Go to a good talkie

187 set.
33. Spend an afternoon visiting a divorce court
34. Spend an afternoon visiting the Ford factory
35. Spend an afternoon visiting an insane asylum
36. Spend an afternoon visiting the United States Senate
133. Take a trip by yourself in a car
170. Be hidden in a closet where you can see and hear a meeting of President Roosevelt's cabinet
171. Be hidden in a closet where you can see and hear a talk between Henry Ford and his son
172. Be hidden in a closet where you can see and hear the private life of a movie star
173. Be hidden in a closet where you can see and hear a criminal being electrocuted

26 Security.
250 set.
240. Spend the evening with your mother

187 set.
95. Be a traffic policeman (in contrast to 94)
124. Travel by train (in contrast to 125)

27 Danger.
187 set.
60. Go from New York to Boston in a small sail-boat
94. Be a regular policeman

125. Travel by airplane
138. Shoot tigers

28 Human Surroundings: Gregariousness.
250 set. 177. Go to a baseball game

187 set. 27. Spend an afternoon at Coney Island
70. Have dinner in a big restaurant
132. Take a trip in a sight-seeing bus
166. Work in the reading room of a library
71. (Reversed). Have dinner in your room alone
167. (Reversed). Work in your own room
177. (Reversed). Spend a day alone in the woods

29 Social intercourse (commonly friendly).
250 set. 80. Going to parties (in teens)
83. Playing around with a group of friends (in teens)
86. Visiting relatives (in teens)
142. Making new acquaintances
145. Going to parties
146. Visiting relatives
154. Being with a group of friends
158. Going to weddings
222. Give a party
237. Go to a card-party
238. Go to a dance
239. Go to a picnic

187 set. 61. Go to a class reunion
62. Go to a club meeting
67. Go on a yacht for a week as a guest of a rich friend
107. Camp out in the Catskills with 10 to 12 friends or acquaintances
168. Be pleasant to strangers
185. Being snowed in for a week in a country hotel with a dozen strangers
186. Having the strangers be friendly to you
187. Living in a social settlement for a week

20 Special Welfare work, talking with old friends, and making new acquaintances.
Sum of ratings among a set of 16

31 Children.
250 set. 65. Taking care of children (in teens)
120. Taking care of children
175. Play with a little child (baby up to 4 years)
176. Play with a child 5 to 10
194. Take care of a baby

187 set. 41. Teach children of six to read
42. Teach children of eight to sing
54. Take pictures of babies
65. Go to visit a day nursery
55. Take pictures of little children

121. Sing a little child to sleep
163. Tell stories to a group of children 8 to 10 years old
180. Take six or eight city children for a day in the country
181. Bring six or eight country children for a day in the city to see the sights

32 Conflict.

250 set.
95. Fighting (in teens)
184. Visit people who owe your firm money and try to collect it.
185. Write letters to people who owe your firm money and get them to pay up
232. Complain to your landlord about something
233. Call a policeman and have him arrest somebody who has cheated you
248. Hearing and adjusting complaints
249. Convincing an opponent that he is wrong

187 set.
49. Make a man give you back your money for an article which he misrepresented
50. Make a professor raise your mark in a course
114. Persuade a person that he or she should resign from candidacy or from office in your favor
123. Tell somebody whom you dislike just what you think of him or her
145. Dispute with a friend over politics, religion, the merits of automobiles, etc.

33 Friendly Rivalry.

250 set.
20. Indoor games (at age 10–14)
54. Card games (in teens)
137. Competitive athletic sports and games
143. Sedentary games like cards and chess

34 Plants and Gardening.

250 set.
55. Working in a garden
89. Growing flowers
90. Growing vegetables
117. Growing vegetables (20 to 25)
118. Growing flowers (20 to 25)
196. See a beautiful garden

187 set.
66. Go to a flower show
80. Picking berries
81. Picking flowers
127. Water a flower garden
128. Weed a flower garden
156. Prune shrubs
164. Water plants in a vegetable garden
165. Weed a vegetable garden
184. Transplanting ferns from the woods to your rock garden

35 Self.

250 set.
144. Your regular job

147. Writing your diary
148. Making plans for the future

For each person for each group we have then a record of so many likes, probable likes, indifferents, probable dislikes and sure dislikes reported. These can be turned into a measure of the person for that group by scoring the indifferents as 0, the likes as plus values of any approved amount (we have used +2 for a sure like and +1 for a probable like), and the dislikes as minus values (we have used −1 and −2). The general fact of the matter is that the items of any group may and probably do represent different 'levels' and also different species at or near the same 'level,' and that the algebraic sum of the credits for the items in any group will be a function of the altitude or level and the width or range of some one 'trait' or tendency with, of course, more or less contamination from the influence of others. For example, suppose that we are concerned with the liking for animals and have honest reports from a person about his likings for n experiences and activities ranging in altitude from playing with a clean and beautiful and affectionate dog to feeding the pigs, or washing a scurvy cur fresh from rolling in a manure pile, and in width so as to include experiences and activities in watching, feeding, tending, nursing, playing with, having as companions, fondling, etc., dogs, cats, rabbits, lambs, horses, cows, canaries, parrots, white mice, etc. We should be able to arrange these experiences and activities according to their attractiveness to some standard population, and then to compute altitude and width scores for any person whose responses to an adequate sample of them were known. Contaminations by other "traits" such as the interest in selling the animals one fed, or in making psychological observations of the animals one tended, or in getting exercise from riding the horse or playing with the dog, would be avoided by proper means.

Let each little square in Fig. 1 represent some one experience or activity, and let all those at Level A be such as 100

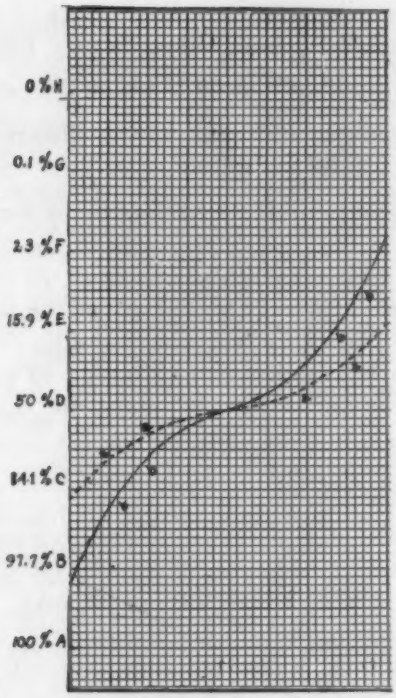


FIG. 1.

per cent of human beings like, all those at Level B be such as 97.7 per cent of them like, all those at Level C be such as 84.1 per cent of them like, all those at Level D be such as 50 per cent of them like, and so on. Then if there is very little specialization, the chief determinant of behavior being the amount of some fairly unitary trait or tendency possessed by the person, any person may be measured by the area below a line such as a a a. If there is much specialization, the line will be more like b b b.

Our scores for the various groups of items are very crude approximations to the measurement of the area below such a line, and so to the average altitude of such a line.

Moreover there probably is a great deal of specialization, so that, for example, a person may be very fond of cats and horses, but not at all fond of cows and birds.

THE RELIABILITIES OF THE MEASURES OF WHATEVER ACCOUNTS FOR A PERSON'S SCORE IN AN ITEM GROUP

I have computed reliability coefficients of various sorts for item groups 7 to 40. I have used half of the 250 set against the other half, both being records taken at the same time, and the 250 set against the 187 set taken months later, with difference in both items and time, and the scores for the 250 and the 187 set together against scores from the Strong Vocational blanks and a special set of 150 questions. The reliabilities are low. There is so much specificity in the likes and dislikes, or so much error in the reports, or both, that the correlation of our scores for item groups 7, 8, 9, 10 . . . 35 with scores from independent sets of items intended to measure identical "traits," will rarely be above .90 and will often be as low as .60, even though there are a dozen items used.

Using the correlation of our scores with those for the Strong + the 150 set where the latter are not from a much smaller number of items, and in such cases making an estimate from the correlations for a part of an item-group against other parts of it, I find the probable correlations of our scores with scores from a comparable independent sampling of activities and experiences to be as follows:

7, .9; 8, .6; 9, .8; 10, .8; 11, .7; 12, .7; 13, .8; 14, .8; 15, .6; 16, .6; 17, .7; 18, low; 19, .9; 21, .8; 22, .5; 23, .7; 24, .6 or .7; 25, .6; 26, low (.3 or .4); 27, low (.3 to .5); 28, low; 29, .5; 30, .8; 31, .8; 32, .3 to .5; 33, low; 34, .7 or .8; 35, low; 36, .7; 37, .8; 38, .5; 39, .5; 40, .6.

Moreover all these self-correlations may be unduly high because of a very strong tendency for persons to vary widely in the proportion of likes to dislikes reported. In a previous article (Journal of Ed. Psy., vol. 26, pp. 497 to 507) we presented evidence showing that in the case of a large group of

adults this tendency seemed to be due to real differences in the total or average enjoyment of a given set of experiences and activities, and much less to constant errors in the interpretation and use of the standards for reporting. But we cannot be sure that such is the case for the 22 persons of the present inquiry. By doubling or quadrupling the number of different items of fact in a group the reliability coefficients can probably be raised to a satisfactory level in many of the groups. But in so far as they are due to a mixture of reality and large constant errors in interpreting and using the standards, the scores will still be misleading as measures of anything. If such constant errors play a very small part our reported likes and dislikes and their intercorrelations are valuable as they stand for both theoretical and practical purposes. If such constant errors play a very large part, we can use their intercorrelations with the general tendency to report too many likes (or dislikes) partialled out, and can adjust individual records for practical purposes more or less satisfactorily in one way or another. But if the reported differences in total enjoyment of any given set of experiences and activities are due to reality and constant errors in proportions from, say, .20–.80 to .80–.20, it will be prudent to change the method.

I turn therefore to a consideration of the ostensible general tendency to like rather than dislike—to enjoy rather than be annoyed by the experiences and activities of our list. The understanding of it is crucial concerning the methodology of measuring personal traits, and the inquiry may disclose a "common factor" or a "principal component" of great interest.

THE TENDENCY TO REPORT A BALANCE OF LIKES OVER DISLIKES (OR DISLIKES OVER LIKES)

If at any one time N individuals report whether they like, dislike, or are indifferent to a miscellany of n experiences and activities, the individual differences in Likes minus Dis-

likes, or in $\frac{\text{Likes minus dislikes}}{\text{Likes plus dislikes}}$ are large and, if n is large, reliable. The distribution of L–D for 31 persons for items 101–250 of the 250 set (*i.e.*, for items 1–250 omitting those reported for childhood or the teens) is shown in Table 1, Column I. The distribution of $\frac{L-D}{L+D}$ is for the same items shown in Table 1, Column III. The correlation of the odds and evens for L–D is .96. The correlation of the odds and evens for $\frac{L-D}{L+D}$ is .91.

Many months later the set of 187 items were rated in the same way as the 250 had been (I am sure that I like it. I think I like it but am not sure. I neither like nor dislike it. I think I dislike it but am not sure. I am sure that I dislike it) by 60 adults, all familiar with this and similar procedures. The distributions of L–D and $\frac{L-D}{L+D}$ are shown in Table 1, Columns II and IV. The correlation of L–D for two random halves of the items was .864. The correlation of $\frac{L-D}{L+D}$ for the two random halves was .84.

A few weeks after these 187 ratings were made, ratings were made for a third set containing 150 new items, by 55 persons from among the 60 who rated the set of 187. The correlation of L–D for items 1–20, 48–75, and 104–131 with L–D for items 21 to 47, 76 to 103, and 132 to 150 is .71. That for $\frac{L-D}{L+D}$ is .75. The correlations using odds and evens would be very much higher. The .71 and .75 give an estimate of the reliability of an L–D measure obtained by taking 26 experiences of taste and smell and a miscellany featuring reading, athletics and art as one half and 35 experiences of sightseeing and of listening plus a miscellany featuring smoking, sedentary games, care of the body. The use of odds and evens would put experiences and activities of each sort nearly

TABLE I

Individual Differences in the Tendency to Like Rather than Dislike According to the Reports: the Relative Frequencies (in Per cents) of Various Values of L–D and of L–D/L+D

L–D	I Frequencies for items 101–250 of the 250 set in 31 persons	II Frequencies for items 1–187 of the 187 set in 60 persons	$\frac{L-D}{L+D}$ (in hundredths)	III Frequencies for items 101–250 of the 250 set in 31 persons	IV Frequencies for items 1–187 of the 187 set in 60 persons
–61 to –70		1.7			
–51 " –60		.....			
–41 " –50	3.2	3.3	–41 to –50		1.7
–31 " –40		3.3	–31 " –40	3.2	1.7
–21 " –30	6.5	3.3	–21 " –30	3.2	5.0
–11 " –20	22.5	3.3	–11 " –20	13.0	5.0
–1 " –10	6.5	6.7	–1 " –10	19.4	8.3
0 " 9	9.7	8.3	0 " 9	9.7	10.0
10 " 19	6.5	5.0	10 " 19	9.7	15.0
20 " 29	19.4	11.7	20 " 29	19.4	18.3
30 " 39	6.5	15.0	30 " 39	6.5	20.0
40 " 49	12.9	16.7	40 " 49	9.7	6.7
50 " 59	6.5	6.7	50 " 59	6.5	6.7
60 " 69		3.3	60 " 69		1.7
70 " 79		3.3			
80 " 89		3.3			
90 " 99		1.7			
100 " 109					
110 " 110		1.7			

equally into each half, and so give too high as a reliability coefficient for a truly random sampling of 75 experiences and activities. The .71 and .75 are too low for a truly random sampling of reports made at the same time. The 150 as a total are not a random sampling, but were chosen deliberately to have a large proportion of sensory experiences and appeals to certain sorts of curiosity plus a miscellany of activities usually considered neither very enjoyable nor very unpleasant, plus also certain items to check certain features of the first list of 250. They are useful here as evidence of the same phenomenon in this very different list of experiences and activities.

The correlation is not so high when the times and items are different as when they are the same and so subject to temporary moods of optimism, and to temporary habits of using the scales, but it is still substantial. The intercorrelations of two random halves of 101–250 of the 250 set, two random halves of the 187 set taken months later, two random halves of 101–200 of the 250 set taken in January with ratings on a scale from +5 to −5, and two random halves of 101–200 of the 250 set taken in September with ratings on the regular scale (sure likes, probably like, indifferent, probably dislike, sure dislike) for the 22 persons specially studied were as shown in Table II.

TABLE II

Intercorrelations of Various L − D (Likes Minus Dislikes) Scores for 22 Adults. A and B Refer to Random Halves. 7 and 8 Differ from 1 and 2 by the Omission of Items 201–250. Decimal Points are Omitted

		1	2	3	4	5	6	7	8
1.	Items 101–250 Sept. A		94	52	57	68	51	93	83
2.	" " " B			54	54	70	62	92	91
3.	" of 187 set Oct. A				87	41	26	62	63
4.	" " " B					47	24	57	54
5.	" 101–200 of 250 set Jan. A						56	66	57
6.	" " " " B							48	36
7.	" " " Sept. A								92
8.	" " " " B								—

The person who reports many likes with any one list on any one day tends strongly to do so with the same list on a day far removed, and also with a list of either similar or very different content on a day far removed.

The reliabilities of L–D for 75, 93 or 50 reports made on the same day by our standard method are .94, .87 and .92. The use of a scale from +5 to −5 reduced the .92 to .56[3].

[3] The ratings of Jan. with items 1–200 of the 250 set were the first made by the group, at a time when they were less accustomed to such self-examination.

When the same items are used, but after months and with a different method, the reliability of L–D for half of items 101–200 is .52. When different items are used and by the same method, but after months, the reliability for half of the items of the 101–250 list or half of the 187 list, is about .54. When the same fifty items plus twenty-five more are used (time and method being identical) the correlation of the fifty with the seventy-five is .90.

L–D for 200 items reported in terms of our scale will then correlate with L–D for 200 other items rated on that day to the extent of .95 or higher, and with 200 other items rated on some day a month or more removed to the extent of .60 or higher. To obtain an accurate measure of an individual's opinions of his likes and dislikes, it is desirable, almost imperative, to gather such opinions on at least four days far apart.

I have used records from five days to get a general measure of L–D, reported as Score 1 in the table of intercorrelations (Table II). Its correlation with a comparable independent measure of L–D should be over .8.

If these individual differences in L–D scores are entirely or chiefly caused by differences in the person's likes and dislikes, we have evidence of the existence and great magnitude of a fundamental tendency to enjoy rather than be neutral or irritated, the possession of which is one of the greatest blessings a person can have, and the extension of which should be one of the great aims of education and eugenics. If they are caused entirely or chiefly by constant errors in interpreting and using the Like-Dislike scale, they provide evidence that measurements of attitudes, interests, and the like by persons' reports are afflicted by systematic errors which raise all intercorrelations and reliability coefficients unduly.

The test is to check the data of persons' reports by data gained from intimate personal knowledge, cross-examination, and experiments in which tendencies to constant error in interpreting and using standards of liking will be absent.

No persons existed who had intimate acquaintance with all the 60 persons of the second experiment, or of the 55 of the third, or of the 31 of the first, or of any large fraction of them. The person with the most knowledge of them was Dr. Lorge, who had spent from 30 to 50 hours a week with all or some of them for nearly a year. Mr. M., one of the group and a very able man, seemed the next most qualified person. This knowledge, however, concerned many of them only under the special circumstances of a group working at learning various things, at making a Semantic word count, and taking multifarious psychological and educational tests.

Dr. Lorge and Mr. M. ranked each of the 31 persons of the first experiment (of whom 22 were included in the second, and 21 in the third) for the strength of their interest in or liking for ideas, things, people, animals, words, reading, bodily exercise, art, music, beauty, outdoor life, responsibility, system, neatness, approval, domination, religion, curiosity, security, danger, sociability, friendship, affection, sex, conflict, achievement, rivalry, money, self, ease, children, talking and clothes, and also for their degree of (I) Cheerfulness, (II) Contentment, and (III) Optimism. For the present I limit the discussion to Dr. Lorge's rankings. His rankings for Cheerfulness, Contentment and Optimism correlate reasonably with each other (.68, .34 and .42); the combined ranking for the three correlates reasonably with the combined ranking for interest in ideas, things, people, animals, . . . money, self, ease, children (.73); and the two random halves of the latter correlate reasonably (.74). Dr. Lorge's sum of ranks for interests in ideas, things, people, etc., correlated with 1 (our combined total measure of L–D) by only .13. His sum of rankings for Cheerfulness, Contentment, and Optimism correlated with 1 by only .34. His sum of ranks and sum of Ch. + Cont. + Opt. correlate .55 with each other. Investigation shows that his ranks were probably subject to certain peculiar halo effects. He ranked the men higher than the women; he ranked the more intelligent much higher than the

less; he ranked the young much higher than the old; he ranked his chief helpers very much higher than the general run.

In my opinion the low correlations between L–D for the persons' reports and Dr. Lorge's rankings are due partly to his lack of acquaintance with the old persons when they were in their teens or at age 20 to 25 and partly to the following: (1) a tendency for certain persons to be much better known to the ranker than others, (2) a tendency for such persons to give him evidence of their positive interests or likes rather than of their dislikes, (3) a tendency for the interest or lack of it shown toward the learning, experiments, and clerical work which occupied the time of the person while they were with Dr. Lorge to spread or halo over the interests in animals, children, housework, etc., and (4) a tendency for mental alertness to spread and halo as well. The validity of such reports as the persons made and the validity of such rankings as a supposedly impartial observer makes of another person's likes and dislikes both evidently need further study.

Unfortunately, by the time that the discrepancy between personal reports and an observer's ratings was known, it was too late to subject the 22 persons to cross-examination concerning their reports, or to further experiments to throw light on the nature and magnitude of their constant errors in using the scale of liking-dislike.

The L–D or $\frac{L-D}{L+D}$ score is at present not divisible into a measured part due to real tendencies to like (or dislike) the various items and another measured part due to constant errors in the use of the Like-Dislike scale. No person familiar with the facts would think it was *all* due to the latter, and I am inclined to the opinion that less than half of it is. But the ambiguity greatly reduces the value of this score, and of all the scores for groups of items. The intercorrelations of the latter are affected with a spurious element to the extent that the constant errors exist, and are affected with a genuine

element of capacity for enjoyment in general to the extent that individual differences in L–D are due to actual likes and dislikes. At present we cannot get rid of the effect of the constant errors without losing the genuine effect also. We can free all the intercorrelations from the influence of both by partialling out Score 1, which is a measure of L–D derived from 487 items[4] and from many items of the Strong Interest Blank.

The intercorrelations appear in Table III, the upper figure in italics being the correlations freed from the influence of Score 1. Scores 26 (Security), 28 (Gregariousness), and 51 (Self) are from so few items as to be very unreliable, but are left in the table with the understanding that their interrelations represent chiefly chance. The row of Means in Table III is in units per item of the score in which a sure like = 2, a probable like = 1, indifference = 0, a probable dislike = −1, and a sure dislike = −2. The row marked S.D./$\sqrt{n}$ gives a rough idea of the intrinsic variation of the group in whatever tendencies are measured by item groups 7, 8, 9, etc., n being the number of items in the group.

I have complete records from only 22 persons, and Table III is limited to these, though in many cases I have had data from 31 that have been studied as a check. The limited number is regrettable. I began with 200 of the 250 items. By the time that the 250 list was ready, many of the subjects were not available. When it was found necessary to have many more items and the 187 set was constructed, the number had dwindled still further. It would have been wiser to have secured many more complete cases, before computing the correlations, so far as this particular inquiry is concerned. But other studies depended upon obtaining guidance from this one. Table III consequently has suggestive rather than demonstrative value; and, in particular, arguments from single correlations will be very unsafe. When several agree in sup-

[4] Items 1–100 of the 250 set which concerned interests in childhood and in the teens were not used in computing 1.

TABLE

Intercorrelations of

	LIKES–DISLIKES	THINGS	PEOPLE	ANIMALS	WORDS	READING	BODILY EXERCISE	ART	MUSIC	BEAUTY	RESPONSIBILITY	DETAIL	SYSTEM	NEATNESS	17 + 18 + 19	APPROVAL	MASTERY	RELIGION
Score	1	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Mean (per item)		–.3	–.3	.3	.4	1.2	.7	.6	1.2	.6	.2	–.7	–.3	–.6	–.6	.2	.1	–.7
S.D./$\sqrt{n}$		3.2	1.7	3.3	1.8	0.9	1.7	3.4	1.8	2.1	2.8	1.3	1.7	2.1	2.5	1.1	1.4	2.6
1		46	45	13	53	03	22	19	41	73	53	45	65	71	71	11	34	17
			–32	*08*	*–32*	*–42*	*40*	*24*	*27*	*–25*	*–02*	*–40*	*–20*	*–19*	*–18*	*14*	*20*	*–12*
7			–05	13	00	–36	44	30	40	18	22	–11	17	27	21	17	32	–02
				–32	*–25*	*12*	*02*	*–39*	*–39*	*–14*	*38*	*56*	*17*	*45*	*51*	*–31*	*13*	*39*
8				–22	04	12	11	–26	–13	24	52	65	41	60	64	–22	26	42
					29	*18*	*–14*	*19*	*26*	*43*	*05*	*–18*	*–21*	*–12*	*–20*	*–24*	*–02*	*–20*
9					31	18	–11	21	29	38	11	–10	–07	01	–05	–22	02	–18
						51	*08*	*38*	*09*	*32*	*–35*	*–32*	*40*	*–20*	*–18*	*08*	*–35*	*–24*
10						44	19	42	28	57	03	–01	60	25	26	13	–10	–11
							00	*09*	*30*	*44*	*–05*	*–08*	*33*	*–05*	*–01*	*04*	*–59*	*23*
11							00	09	29	32	–03	–06	27	–02	01	04	–55	23
								05	*16*	*00*	*04*	*–12*	*30*	*22*	*18*	*26*	*–06*	*–04*
12								09	23	17	15	00	37	31	28	37	02	00
									38	*20*	*01*	*–49*	*04*	*–26*	*–32*	*16*	*–40*	*–10*
13									42	27	11	–35	09	–05	–09	18	–31	–07
										19	*23*	*–06*	*28*	*15*	*15*	*31*	*–58*	*28*
14										42	39	13	46	39	39	33	–37	32
											38	*–03*	*04*	*–01*	*–04*	*–42*	*–62*	*19*
15											61	31	50	52	–50	–20	–15	25
												35	*21*	*47*	*48*	*–12*	*–02*	*27*
16												51	48	66	66	–04	16	32
													25	*73*	*83*	*–31*	*–04*	*48*
17													46	78	84	–23	12	50
														49	*60*	*14*	*–35*	*08*
18														71	77	18	–03	17
															96	*–04*	*–11*	*52*
19															98	05	17	48
																–10	*–15*	*48*
20																01	14	46
																	–10	*01*
21																	–05	03
																		–49
22																		–40
23																		
24																		
25																		
26																		
27																		
28																		
29																		
30																		
31																		
32																		
33																		
34																		
35																		
36																		
37																		
38																		
39																		
40																		
41																		
42																		

TABLE III

rcorrelations of Item Group Scores

MASTERY	RELIGION	DISPLAY	NOVELTY	SECURITY	DANGER	HUMAN SURROUNDINGS	SOCIAL INT.	WELFARE WORK, ETC.	CHILDREN	CONFLICT	FRIENDLY RIVALRY	PLANTS AND GARDENING	SELF	REALITY	TALKING	IMAG., FICTION	INDULGENCE	MORAL	SED. GAMES	DANCING	INTELLIGENCE TEST SCORE
22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	CAVD
.1	−.7	−.2	.4	.5	−.2	.1	.3	—	−.1	−.6	.1	.4	.7								
1.4	2.6	2.3	1.2	2.0	1.1	1.8	2.3	—	3.0	1.7	1.2	2.2	1.5								
34	17	67	35	−01	65	−10	37	20	56	33	21	45	22	65	54	52	26	70	48	05	03
20	*−12*	*−10*	*−01*	*17*	*−34*	*−51*	*−75*	*−53*	*−74*	*−60*	*−14*	*−15*	*−27*	*−09*	*−49*	*−32*	*−13*	*−15*			
32	−02	24	15	15	08	−40	−45	−37	−29	−36	−02	08	−13	24	−12	−01	01	23	60	−48	16
13	*39*	*11*	*−16*	*56*	*12*	*51*	*53*	*35*	*−01*	*22*	*17*	*−30*	*35*	*16*	*54*	*13*	*01*	*30*			
26	42	37	−02	50	37	41	60	39	24	34	25	−04	40	40	65	33	12	49	23	29	−36
−02	*−20*	*−38*	*02*	*−27*	*−34*	*02*	*−47*	*−37*	*42*	*−20*	*−08*	*30*	*−01*	*10*	*−28*	*34*	*02*	*06*			
02	−18	−19	06	−27	−17	00	−38	−33	42	−14	−05	32	02	16	−17	35	05	14	10	−07	−39
−35	*−24*	*17*	*32*	*−21*	*−09*	*−25*	*−17*	*04*	*24*	*−21*	*−52*	*49*	*26*	*−18*	*09*	*56*	*−38*	*−34*			
−10	−11	46	44	−19	28	−27	07	13	46	01	−32	61	33	23	35	68	−17	16	03	−07	05
−59	*23*	*23*	*06*	*−03*	*−32*	*09*	*15*	*36*	*34*	*−37*	*−11*	*−07*	*03*	*−16*	*34*	*67*	*−42*	*−02*			
−55	23	19	07	−03	−23	08	15	36	30	−34	−10	−05	03	−10	30	58	−40	00	−11	35	−15
−06	*−04*	*65*	*17*	*13*	*44*	*−29*	*13*	*−05*	*−07*	*21*	*22*	*33*	*−15*	*61*	*26*	*17*	*40*	*−18*			
02	00	62	23	13	47	−26	20	−01	07	27	25	39	−10	60	34	26	44	03	07	−12	31
−40	*−10*	*06*	*26*	*−25*	*10*	*−44*	*−44*	*−33*	*23*	*−17*	*−57*	*25*	*−14*	*−19*	*−12*	*27*	*−45*	*−29*			
−31	−07	17	30	−25	20	−45	−33	−28	29	−09	−51	30	−09	−02	00	32	−38	−07	−26	01	−19
−58	*28*	*12*	*00*	*12*	*−30*	*−22*	*−43*	*05*	*10*	*−20*	*−24*	*−09*	*−14*	*09*	*−06*	*20*	*−36*	*04*			
−37	32	36	14	10	05	−25	−22	12	30	−04	−13	11	−03	33	18	37	−21	31	02	05	−11
−62	*19*	*−13*	*−27*	*00*	*12*	*15*	*−05*	*27*	*49*	*−41*	*−17*	*03*	*07*	*04*	*26*	*22*	*26*	*11*			
−15	25	42	09	00	54	00	24	33	69	−02	04	35	21	50	54	51	36	57	33	33	−23
−02	*27*	*−38*	*−06*	*39*	*08*	*48*	*03*	*33*	*21*	*−15*	*−03*	*−22*	*11*	*63*	*10*	*−24*	*00*	*−01*			
16	32	11	14	32	39	35	22	38	44	06	09	07	20	75	36	10	14	37	22	32	−49
−04	*48*	*11*	*−19*	*58*	*−36*	*55*	*36*	*45*	*10*	*13*	*−01*	*−43*	*01*	*00*	*28*	*−14*	*−01*	*55*			
12	50	37	−01	51	04	44	47	48	32	26	09	−15	15	31	48	13	12	67	14	34	−32
−35	*08*	*24*	*05*	*42*	*−17*	*12*	*09*	*47*	*23*	*−01*	*−20*	*−03*	*12*	*25*	*40*	*19*	*−30*	*18*			
−03	17	56	26	31	31	03	29	47	50	20	−01	27	22	56	60	45	−06	55	13	14	04
−11	*52*	*15*	*00*	*71*	*−28*	*36*	*16*	*21*	*02*	*−16*	*−10*	*−10*	*31*	*29*	*29*	*−07*	*−09*	*62*			
17	48	55	25	49	32	17	37	28	41	13	08	26	36	62	56	33	13	81	23	12	−28
−15	*48*	*18*	*−05*	*74*	*−45*	*40*	*20*	*43*	*−05*	*−01*	*−09*	*−22*	*25*	*26*	*37*	*−10*	*−12*	*58*			
14	46	57	21	51	21	21	39	44	42	23	09	18	34	60	60	31	10	79	21	31	−24
−10	*01*	*24*	*58*	*−11*	*07*	*−15*	*−35*	*03*	*−29*	*−46*	*−19*	*26*	*13*	*13*	*−27*	*07*	*−12*	*−04*			
−05	03	25	58	−11	13	−16	−28	05	−18	−40	−16	28	15	17	−16	12	−09	05	−05	−21	30
	−49	*−78*	*13*	*09*	*35*	*09*	*−21*	*−36*	*−58*	*33*	*21*	*31*	*24*	*10*	*−59*	*−63*	*41*	*12*			
	−40	−32	23	09	47	05	−06	−27	−27	40	28	41	30	29	−29	−33	46	32	41	−46	00
		47	*−26*	*53*	*−38*	*17*	*54*	*28*	*−01*	*−24*	*−18*	*−33*	*25*	*02*	*50*	*32*	*−28*	*60*			
		46	−18	52	−17	15	56	31	09	−16	−14	−24	28	13	51	36	−22	54	−03	34	−19
			−09	*30*	*−07*	*−13*	*31*	*09*	*09*	*−13*	*−25*	*−03*	*−06*	*27*	*83*	*48*	*−35*	*53*			
			17	22	39	−17	46	31	32	16	04	23	18	42	78	56	−02	54	08	13	23
				−33	*08*	*−20*	*−14*	*−01*	*−15*	*−38*	*−46*	*42*	*11*	*−05*	*−50*	*02*	*−20*	*−11*			
				−32	28	−22	01	06	08	−22	−34	51	17	19	−21	19	−09	17	−07	−18	04
					02	*19*	*19*	*44*	*−27*	*10*	*15*	*04*	*39*	*51*	*57*	*−09*	*10*	*31*			
					01	19	18	43	−23	09	14	03	38	38	48	−08	10	21	05	10	−30
						−33	*−32*	*−04*	*−47*	*18*	*09*	*31*	*18*	*32*	*−04*	*−41*	*22*	*−41*			
						−18	01	10	06	35	20	49	28	61	33	07	33	23	33	20	04
							38	*−10*	*46*	*−09*	*−01*	*−77*	*−03*	*28*	*−05*	*05*	*−26*	*12*			
							32	−12	32	−12	−03	−73	05	14	−10	−02	−28	01	38	01	−34
								22	*05*	*−02*	*07*	*−37*	*06*	*−13*	*44*	*21*	*11*	*39*			
								27	24	10	14	−14	14	15	54	36	19	52	20	28	18
									−13	*−32*	*06*	*−20*	*27*	*16*	*48*	*−18*	*−01*	*15*			
									01	−23	10	−09	31	25	50	−05	04	25	17	27	−11
										−12	*−25*	*−31*	*−33*	*−16*	*14*	*33*	*−35*	*33*			
										09	−09	02	−15	26	40	52	−13	58	−04	42	−18
											25	*−24*	*−35*	*−10*	*02*	*−24*	*24*	*−30*			
											31	−06	−25	15	20	−02	31	04	18	−30	21
												−11	*−28*	*15*	*13*	*−37*	*57*	*34*			
												00	−22	25	22	−19	60	38	71	−04	19
													35	*29*	*−22*	*−03*	*09*	*−33*			
													40	49	07	21	19	10	33	−43	00
														26	*12*	*22*	*−10*	*07*			
														33	22	30	−03	20	14	−27	31
															10	*24*	*30*	*26*			
															42	18	39	60	25	−11	10
																30	*−13*	*22*			
																50	04	51	19	36	−08
																	05	*07*			
																	18	41	08	23	−13
																		−12			
																		10	53	−09	07
																			20	33	−17
																				−18	09
																					−76

porting a certain hypothesis, some confidence may be felt, and characteristics of the table as a whole, though conceivably due to peculiarities of the small population concerned, are probably worth very serious consideration.

I shall present in order some facts and hypotheses concerning factors determining the intercorrelations, the specificity of interests, and certain unsuspected relations among interests.

GENERAL LIKE

In an article presenting reported degrees of like and dislike for 16 sorts of activity or experience in the case of some two hundred adults, I showed that the total or average amount of liking varied widely among individuals. The same is true of the individuals of the present study, unless we assume an unbelievably great amount of variation in the interpretation and use of the scale used. A large amount of such there undoubtedly is. But among the 22 persons with complete records the Likes minus Dislikes difference varies too much for me to be willing to explain it all by such constant errors. Moreover the correlations of the separate items with the Likes – Dislikes difference vary too much. If Likes – Dislikes were entirely a resultant of constant errors of standards, these intercorrelations should vary by approximately their unreliability, and in a normal surface of frequency. Their variation is greater than that by a fourth and its form is much more like a rectangle with a tail at the low end. Also the high correlations with Likes – Dislikes were for items where persons should not be specially liable to constant errors in using the standards. They were for beauty, system, neatness, sum of detail, system and neatness, display, and danger. Also Likes–Dislikes correlates zero with intelligence, whereas if it were due to constant errors it should show a substantial inverse correlation therewith.

SEX

Sex in this group is a potent factor, as appears from the mean scores shown in Table IV. The reliabilities of the dif-

TABLE IV

Differences of Means of Men and Women (M–W) as Multipliers of the Standard Deviation for Men and Women Taken Together

1.	Likes–Dislikes (Composite)	– .24
2.	Likes–Dislikes (250 + 187 that are in 7–35 of this list)	– .45
3.	Likes–Dislikes (150 + Strong)	– .17
4.	Likes–Dislikes (250 + 187 that are not in 7–35 of this list)	+ .04
5.	Likes–Dislikes as estimated by L.	+ .33
6.	Composite of ratings by L for Cheerfulness, Contentment and Opinion	+ .35
7.	Things	+ .70
8.	People	– .07
9.	Animals	– .35
10.	Words	– .73
11.	Reading	– .84
12.	Bodily exercise	+ .42
13.	Art	– .69
14.	Music	– .51
15.	Beauty	–1.28
16.	Responsibility	– .78
17.	Detail	– .25
18.	System	– .61
19.	Neatness	– .58
20.	17 + 18 + 19	– .50
21.	Approval	+ .01
22.	Mastery of persons	+ .99
23.	Religion	– .70
24.	Display: public attention	– .16
25.	Novelty; curiosity	– .18
26.	Security	+ .06
27.	Danger	+ .73
28.	Human surroundings: gregariousness	– .42
29.	Social intercourse	– .23
30.	Ratings for welfare; talking with old friends and making new acquaintances	– .51
31.	Children	–1.03
32.	Conflict	+ .99
33.	Friendly rivalry	+ .96
34.	Gardening; plants	+ .05
35.	Self	– .39
36.	Reality	– .10
37.	Talk	– .50
38.	Imagination; fiction	– .68
34.	Gardening; plants	+ .05
40.	Moral	– .67
41.	Sedentary games	+ .64
42.	Dancing	–1.38

ferences are too low to justify decisions about sex and particular interests in adults, but from general observations and certain other studies we may be sure that it is a substantial factor.

DETAIL, SYSTEM, NEATNESS

Whatever is measured by the reports listed under any one of these three is much the same as what is measured under any other of the three. Their intercorrelations, .46, .78 and .71, are nearly as high as their reliability coefficients permit. The intercorrelations after partialling out 1 are .25, .73, and .46.

Whatever links 17, 18 and 19 is in our population linked to 'femaleness,' 'people,' 'bodily exercise,' 'security,' 'social intercourse' and 'moral,' and opposed to 'maleness,' 'animals,' 'words,' 'beauty,' 'dancing' and intellectual ability. This is a rather strange pattern, and may be due to the accidental constitution of these 22 persons concerned. But I think there is method in the apparent madness.

OTHER LINKED INTERESTS

Animal, Beauty and Child show intercorrelations of .38, .42 and .69, and, after 1 is partialled out, of .43, .42 and .49. Two of these are markedly feminine interests, and the animal of our items is a rather gentle and domestic composite. But something more than sex links these. All three are linked to Art (partials of .19, .20 and .23), Music (partials of .26, .19 and .10), Words (partials of .29, .32 and .24), Reading (partials of .18, .44 and .34) and Imagination or Fiction (partials of .34, .22 and .33). They are not closely linked to the 17, 18, 19 group (partials of –.18, –.21, –.12, –.06, .28, .15, .10, .22, and .02) nor to Religion (partials of – .20, .19 and – .01), nor to Security (partials – .27, .01 and – .27.).

Words, Reading, Art, Music and Beauty have correlations (lower row) and partial correlations (upper row) as shown in Table V.

TABLE V
Interrelations of Item Groups 10, 11, 13, 14, and 15

	READING	ART	MUSIC	BEAUTY
Words	51 44	38 42	09 28	32 57
Reading		09 09	30 29	44 32
Art			38 42	20 27
Music				19 42

The average inter-correlation with 1 partialled out is .32 for Words, .34 for Reading, .26 for Art, .24 for Music, and .29 for Beauty. These groups are allied to femininenes, and opposed to People and Social Intercourse. They are opposed to Mastery, Conflict, and Rivalry even more than their femininenes would account for.

Mastery over Persons, Danger, Conflict and Friendly Rivalry are all linked to Masculinity, and among themselves. With 1 partialled out, their intercorrelations are .35, .37, .22, .18, .09 and .25.

People, Social Intercourse, and Welfare-Work + Talking, etc., have partial r's of .53, .35, and .22. Gregariousness has partial r's with them of .51, .38 and .10 respectively.

Other affiliations can be traced in Table III. The unreliabilities due to the small number of cases makes the conclusions from the observed affiliations hazardous. But when the partial intercorrelations of several average far above or below the +.05 which is the median of Table III, they deserve consideration.

THE SPECIALIZATION OF INTERESTS

Interests are allied, but they are rarely allied very closely. Art, Music, and Beauty, for example, are correlated to the extent of .42, .27 and .43; and if freed from attenuation these

correlations would perhaps be .7, .5, and .7. People correlates with Children only .24, which correction might raise to .4 or .5. The highest correlations with intelligence are .31 for Bodily Exercise, .30 for Approval, and .31 for Self. The lowest are −.76 for Dancing, −.49 for Responsibility, and −.39 for Animals.

THE ALTERNATION OR SUBSTITUTION OF INTERESTS

Popular psychology often assumes that there is a certain quantum of a certain favorable attitude in a certain person and that if this is directed to one sort of object it must be at the cost of some other sort or sorts of objects. There should then be (if the measures are adequate) a perfect negative correlation between the two in a group of persons in whom the quantum was the same for all. The total flow of kindly regard for the weak and dependent might thus be divided between children and animals. The total flow of diffuse interest might thus be divided between things, persons and ideas. The total flow of interest in action might thus be divided between bodily exercise, sedentary games, useful productive labor, and fidgeting.

We cannot test this doctrine satisfactorily from our records, but in general the negative correlations after partialling out 1 are much nearer zero than one would expect from this doctrine (Things and People, −.32; Things and Words, −.32; Things and Reading, −.40; Reality and Fiction, −.24; etc., etc.).

SOME UNSUSPECTED RELATIONS

Indulgence and Moral represent fairly well what the terms conventionally imply. They are not antagonistic, but show a slight positive r (.10); and even when 1 is partialled out, the drop is only to −.12. By our reports, the person who likes the decent, useful, and virtuous likes the world, the flesh, and the devil little or no less than others. If a wider sampling substantiates this fact, the consequences to the theory and practice of morals are important.

TABLE V

Interrelations of Item Groups 10, 11, 13, 14, and 15

	READING	ART	MUSIC	BEAUTY
Words	51 44	38 42	09 28	32 57
Reading		09 09	30 29	44 32
Art			38 42	20 27
Music				19 42

The average inter-correlation with 1 partialled out is .32 for Words, .34 for Reading, .26 for Art, .24 for Music, and .29 for Beauty. These groups are allied to femininess, and opposed to People and Social Intercourse. They are opposed to Mastery, Conflict, and Rivalry even more than their femininess would account for.

Mastery over Persons, Danger, Conflict and Friendly Rivalry are all linked to Masculinity, and among themselves. With 1 partialled out, their intercorrelations are .35, .37, .22, .18, .09 and .25.

People, Social Intercourse, and Welfare-Work + Talking, etc., have partial r's of .53, .35, and .22. Gregariousness has partial r's with them of .51, .38 and .10 respectively.

Other affiliations can be traced in Table III. The unreliabilities due to the small number of cases makes the conclusions from the observed affiliations hazardous. But when the partial intercorrelations of several average far above or below the +.05 which is the median of Table III, they deserve consideration.

THE SPECIALIZATION OF INTERESTS

Interests are allied, but they are rarely allied very closely. Art, Music, and Beauty, for example, are correlated to the extent of .42, .27 and .43; and if freed from attenuation these

correlations would perhaps be .7, .5, and .7. People correlates with Children only .24, which correction might raise to .4 or .5. The highest correlations with intelligence are .31 for Bodily Exercise, .30 for Approval, and .31 for Self. The lowest are −.76 for Dancing, −.49 for Responsibility, and −.39 for Animals.

THE ALTERNATION OR SUBSTITUTION OF INTERESTS

Popular psychology often assumes that there is a certain quantum of a certain favorable attitude in a certain person and that if this is directed to one sort of object it must be at the cost of some other sort or sorts of objects. There should then be (if the measures are adequate) a perfect negative correlation between the two in a group of persons in whom the quantum was the same for all. The total flow of kindly regard for the weak and dependent might thus be divided between children and animals. The total flow of diffuse interest might thus be divided between things, persons and ideas. The total flow of interest in action might thus be divided between bodily exercise, sedentary games, useful productive labor, and fidgeting.

We cannot test this doctrine satisfactorily from our records, but in general the negative correlations after partialling out 1 are much nearer zero than one would expect from this doctrine (Things and People, −.32; Things and Words, −.32; Things and Reading, −.40; Reality and Fiction, −.24; etc., etc.).

SOME UNSUSPECTED RELATIONS

Indulgence and Moral represent fairly well what the terms conventionally imply. They are not antagonistic, but show a slight positive r (.10); and even when 1 is partialled out, the drop is only to −.12. By our reports, the person who likes the decent, useful, and virtuous likes the world, the flesh, and the devil little or no less than others. If a wider sampling substantiates this fact, the consequences to the theory and practice of morals are important.

The correlations with Intelligence score make an interesting list. Table VI shows them arranged according to their mag-

TABLE VI

*Correlations with Intelligence Score**

Dancing	−76	Gardening; plants	00
Responsibility	−49	Mastery	00
Animals	−39	Gen. Like	03
People	−36	Danger	04
Gregariousness	−34	Novelty	04
Detail	−32	System	04
Security	−30	Words	05
Neatness	−28	Indulgence	07
Beauty	−23	Sedentary games	09
Art	−19	Reality	10
Religion	−19	Social intercourse	18
Children	−18	Friendly rivalry	19
Moral	−17	Conflict	21
Reading	−15	Display	23
Imagination; fiction	−13	Approval	30
Music	−11	Bodily exercise	31
Welfare work, talk, etc.	−11	Self	31
Talk	−08		

* The Intelligence Score has a reliability near .98.

nitude. They are complicated by the fact that in the 22 persons intelligence is positively correlated with sex, and by the general unreliability of Items 7 to 43. I call attention especially to the antagonism of Intelligence to Responsibility and People, and its affiliation with Approval. The influence of sex is not adequate to account for these; and I surmise that they are not chance results of our scanty samplings, and suggest that the frequent failure of intellectuals to use their power in politics, business, and human affairs generally is due in part to their dislike of people and responsibility and their sensitiveness to criticism and blame.

Talking is a rather definite and specific activity and it is not hard to conceive how individuals could by nature or by training or by both differ in respect of enjoyment of it.

Some of its correlations, after partialling out 1, are as follows: Display, .67; People, .54; Religion, .50; Welfare Work and Talk with Old Friends and Making New Acquaintances, .48; Social Intercourse, .44; Words, .09; Music, −.06; Gardening and plants, −.22; Animals, −.28; Things, −.49; Novelty or Curiosity, −.50; Mastery, −.59.

Some of the items in Talk occurred also in Groups 7 to 35 and our so-called 'People' and 'Display or Public Attention' groups included large dosages of display by speech. But, even after generous allowance for these facts and for the moderate correlation of Talk with Femininity, the liking for Talk as such seems a potent factor. It seems to have little or nothing to do with the liking for words as such, or with the liking for sounds as such.

THE ROLE OF FACULTY COUNSELING IN SCHOLASTIC MOTIVATION*

E. G. WILLIAMSON
University of Minnesota

THE purpose of student personnel work has been stated as "the satisfactory educational adjustment of each student" and "the individualization of mass education to the end that each student will eventually make an adequate adjustment to the requirements of life"—"including effective socialization, living according to the principles of physical and mental hygiene, developing one's potential intellectual interests and powers to the utmost, developing a high type of moral conduct, acquiring vocational and professional knowledge and skill to insure adequate service to society, and learning how to utilize leisure time profitably"—"the educational process must modify, develop, and coordinate the capacities, interests and opportunities of each individual student" (1).

The University of Minnesota agencies that exist to achieve these objectives are numerous. In this discussion, however, we are concerned only with one of these agencies, namely the faculty counselors in the Arts College, who act as a personal medium for contact between the individual advisees and the rest of the University in many of the above adjustments. The specific problem studied was the adjustment of freshmen to the scholastic requirements of the college. The criterion of the effectiveness of counseling was scholarship achievement of advisees as compared with non-advised students, the two groups being matched in ability and high school scholarship.

Langlie and Paterson studied the effect on scholarship of faculty counseling in the College of Engineering and Architecture at the University of Minnesota in the academic year

* This project was presented as part of the author's doctoral thesis in psychology and was conducted under the direction of Professor D. G. Paterson.

1925–1926 (2). Twelve faculty counselors interviewed 60 probationers between mid-quarter examinations and the end of the fall quarter. These probationers were compared with 30 students who made "superior" scholastic records in the first six weeks of the fall quarter with regard to academic abilities, study habits, and personality traits. The probationers were lower than the "superior" students in rankings on Iowa aptitude tests, college aptitude test, and on "ratings of nervous strain." After the interviewing, the probationers were divided at random into two equal groups, one of which served as a control group and the members of the other group were assigned to faculty counselors for "intensive follow-up guidance and assistance." These two groups were equated on the basis of information similar to that used in this present study. This experiment showed that the advised students did no better scholastic work than the non-advised probationers.

Johnston has stated that students with a marked discrepancy between their high school scholarship and college aptitude rating could expect to have difficulty with college work unless they had special attention (3). With this fact in mind, steps were taken to aid and, if possible, save some of these discrepancy students. In the fall of 1926–27, seventy freshmen in the College of Science, Literature, and the Arts who had a discrepancy of 30 or more points between the percentile ranks of the Minnesota college aptitude test and high school scholarship were assigned to faculty counselors. Nine of these seventy students were not advised because they failed to keep appointments with the counselors; therefore, this analysis is based upon sixty-one students who were advised and sixty-one matched not-advised students who were not assigned to faculty counselors. These sixty-one students were paired with sixty-one control students in college aptitude test rank and high school scholarship rank.

Both the control and the experimental groups were divided into two separate groups, one group containing those students whose percentile rank in the college aptitude test was higher

than their rank in high school scholarship; and a second group, the members of which ranked higher in high school scholarship than in the college aptitude test. The reason for such a dichotomy is that it has been discovered both in research studies and in actual counseling that these two types of students have different scholastic problems. The student who has high academic ability but low scholastic achievement is more apt to have inefficient study habits and a disinterested attitude toward studies. The student who has lower academic ability than scholastic achievement is more likely to be an ambitious student who works diligently and may achieve beyond the normal expectancy based on his rank in aptitude tests. There are more boys than girls with high test ranks and low high school scholarship; but the number of girls exceeds the number of boys with the opposite kind of discrepancy. This sex division on the basis of the kind of discrepancy is in line with the finding of Paterson and Langlie that girls with low academic ability tend to get higher grades than do boys with comparable ability (4).

Each counselor participating in this experiment made an intensive study of the student's life history, giving particular attention to factors that might interfere with his maximum scholastic performance, and removing these obstacles in so far as it was possible. The aid given to the students involved the improvement of study methods, budgeting time, overcoming tendencies to worry, and the encouragement of more social contacts when students seemed to be under-socialized. Each adviser saw his advisees many times during the academic year. Information concerning the student's intelligence, high school scholarship rank, family economic and educational background and other data was given to all counselors.

The number of students involved in the 1926–27 experiment is small but, within the limitations of experimental conditions, it is apparent that advised students did no better scholastic work than did not-advised freshmen, except in the spring quarter. Table I gives the comparison in median honor point

ratio[1] and percentage overlapping of the advised and not-advised students classified as to kind of discrepancy, and for all advised and not-advised students. The students who were high in aptitude and low in high school scholarship have practically the same average scholarship as have the students who were not advised by faculty counselors. This same trend is present for those students low in aptitude and high in high school scholarship. When both groups of freshmen are combined, there is no significant difference in average scholarship. Furthermore, there is no significant difference between the advised and not-advised groups regarding the number of students who cancelled from the University. Jones (5), Lemon (6), and Holaday (7) found that freshmen who had been counseled remained in college longer than did non-advised students.

The average scholarship, however, of both advised and not-advised students increased in the winter and spring quarters. The significant increase in the scholarship of the advised group during the spring quarter may be explained in terms of the advice given by the counselors during the fall and winter quarters. Perhaps it has taken these students two quarters to profit from the counselor's advice with regard to study habits. Another interpretation of this increase in scholarship is found in the fact that twelve of the advised and ten of the not-advised group cancelled their registration, did not register at all during the quarter, or were dropped for low scholarship. These not-advised students who were not in residence during the spring quarter had a slightly higher fall quarter scholarship, -0.10, than did the advised students who dropped out before or during the spring quarter, -0.27. In other words, the significantly higher spring quarter scholarship of the advised students may be due to the fact that the counselors had persuaded the students low in scholarship to leave college. Such

[1] Honor point ratio is found by dividing the total number of honor points earned by the total number of credit hours carried. A receives three honor points, B two, C one, D none, and F minus one honor point.

a factor may explain Holaday's results. Because there is no corresponding influence at work with the not-advised students, except what is common to all students, fewer of these low stu-

TABLE I

Comparing the College Scholarship of Advised with Not-Advised 1926–1927 Freshmen

	GROUP A		GROUP B		ALL STUDENTS	
	High Aptitude and Low High School Scholarship		Low Aptitude and High High School Scholarship		Advised	Not-Advised
	Advised	Not-Advised	Advised	Not-Advised		
Median Fall Quarter Honor Point Ratio	0.50	0.58	1.0	0.93	0.78	0.74
Per Cent of One Group Reaching or Exceeding Median of the Other Group	49	59	55	47	53	52
No. Cancellations	.....	.....	.....	.....	.....	.....
Median Winter Quarter Honor Point Ratio	0.83	0.90	0.85	1.10	0.80	1.05
Per Cent of One Group Reaching or Exceeding Median of the Other Group	45	47	64	64	38	59
No. Cancellations	5	5	5	5	10	10
Median Spring Quarter Honor Point Ratio	1.0	0.90	1.21	1.04	1.11	1.01
Per Cent of One Group Reaching or Exceeding Median of the Other Group	57	45	71	24	61	35
No. Cancellations	4	4	8	6	12	10

dents left college and consequently their median spring quarter scholarship is lower than that of the advised group.

The median honor point ratio for the three quarters of the freshman year for the A advised group was 0.50, 0.83, 1.0, as compared with 0.58, 0.90, 0.90 for the A not-advised group. The median honor point ratio for the B advised group was 1.0, 0.85, 1.21 and for the B not-advised group 0.93, 1.10, 1.04 for the three quarters. This increase in spring quarter scholarship is present regardless of the direction of the discrepancy between college aptitude test and high school scholarship.

A similar experiment was conducted during the academic year 1927–1928 by selecting one hundred fifty students who had a discrepancy of ten or more points between college aptitude and high school scholarship percentiles. These students were divided into two groups by pairing individual cases on the basis of college aptitude, high school scholarship, and college aptitude rating. The members of one group were assigned to faculty counselors and the members of the other were used as controls. The counselors used a standard interview form of procedure in their study of advisees, which yielded comparable data on all such cases. The information obtained included number of hours of sleep, eating habits, number of hours, plan and method of study, reasons for selection of course, hours given to activities and outside work each week, a judgment of the emotional and nervous balance of the individual, and an estimate of his attitude toward instructors, courses, hours of class work, family, financial situation, and social life.

One hundred fifty students not assigned to faculty counselors were asked to fill out a questionnaire giving the same information as collected by counselors in interviews with the experimental group of advisees. Although these two methods of collecting this information may have produced differences in the validity of the data, yet this procedure seemed the only possible way of insuring that the control group would not be advised by counselors.

The faculty counselors worked with these students from the time they were assigned in October through the remainder of the fall quarter and in most cases continued seeing them throughout the year. Presumably, since our cases are well matched, if faculty counselors have any favorable influence on scholarship, it would be revealed with these discrepancy advisees. In order to answer this question of effect of counseling on scholarship, we have analyzed our data in the following manner:

1. What differences are there in average scholarship and percentage overlapping between the advised and not-advised groups?

2. What influence has the amount of discrepancy between college aptitude test and high school scholarship upon college scholarship?

Table II gives the median scholarship, percentage overlapping, and number of students cancelling from college of the advised and not-advised freshmen for all three quarters of 1927–28. There are no significant differences in average scholarship or percentage overlapping between the advised and not-advised students except for the spring quarter. A difference appears in the case of those students high in aptitude but low in high school scholarship, the median honor point ratio for the advised freshmen being 1.10 and 0.32 for the not-advised. Ninety per cent of the advised group equal or exceed the median of the not-advised. This difference in average scholarship for the spring quarter may be due to the fact that records were available for only 47 pairs of students. The number of withdrawals from college because of low grades or other factors is about the same for both advised and not-advised students for the winter and spring quarters. But in the spring quarter 7 of the advised students of Group B cancelled as compared with 13 of the not-advised students. One explanation of this difference is that counselors may have helped the low ability students sufficiently for them to get passing grades. Since such a difference in favor of the coun-

TABLE II

Comparison of College Scholarship of Advised with Not-Advised 1927–28 Freshmen

	GROUP A		GROUP B		ALL STUDENTS	
	High Aptitude and Low High School Scholarship		Low Aptitude and High High School Scholarship		Advised	Not-Advised
	Advised	Not-Advised	Advised	Not-Advised		
Median Fall Quarter Honor Point Ratio	0.80	1.04	1.15	1.21	1.13	1.14
Per Cent of One Group Reaching or Exceeding Median of the Other Group	.44	.56	.43	.54	.49	.49
No. Cancellations	—	—	—	1	—	1
Median Winter Quarter Honor Point Ratio	0.78	0.76	1.02	1.28	1.15	1.21
Per Cent of One Group Reaching or Exceeding Median of the Other Group	52	48	48	43	44	55
No. Cancellations	2	2	5	7	7	9
Median Spring Quarter Honor Point Ratio	1.10	0.32	1.14	1.28	1.19	1.18
Per Cent of One Group Reaching or Exceeding Median of the Other Group	90	14	38	56	47	49
No. Cancellations	5	7	7	13	12	20

selors is present only in this one situation, it seems more reasonable that this greater number of cancellations among the not-advised cases is due to other causes.

If whatever difference in scholarship counseling may produce does not appear year after year with any degree of consistency, then we are justified in concluding that chance factors in the elimination of students of different levels of achievement is responsible for such differences as do appear. Therefore we conclude that within the limitations of these experimental conditions, counseling, such as practiced in this experiment, has no appreciable effect upon the scholarship of freshmen in the College of Science, Literature, and the Arts.

Does the amount of discrepancy between college aptitude and high school scholarship influence the effectiveness of advising? For instance, do the 43 advisees with a discrepancy of 10 to 39 points between college aptitude test and high school scholarship percentiles profit more than do the 15 advisees with a discrepancy of 40 to 69 points? If we compare the median scholarship of those advisees who have a discrepancy between 10 and 39 points with those whose discrepancy is between 40 and 69 points, we find no significant difference. It was possible to make this comparison only where the aptitude was lower than high school scholarship because of the limited number of cases. There were too few cases of the discrepancy in which the high school scholarship was lower than the college aptitude test to make a similar comparison as that given in Table III. The direction of the discrepancy was disregarded in this comparison.

TABLE III

Comparing the College Scholarship of Students with Different Amounts of Discrepancy Between Aptitude and Achievement

	10 TO 39 POINTS		40 TO 69 POINTS	
	Advised	Not-Advised	Advised	Not-Advised
Median Fall Quarter Honor Point Ratio	1.12	1.2	1.1	1.18

But it is possible that the effectiveness of counseling is dependent upon the level of intelligence or the level of performance in high school scholarship. Our data were analyzed to

discover if such were the case. Using the mean honor point ratio we have compared the scholarship of our advisees divided into lower, middle, and upper one-thirds on the basis of college aptitude, high school scholarship and college aptitude rating.

In the following table we find an indication that there is most improvement in the case of advisees in the lowest one-third in college aptitude, middle one-third in high school scholarship, and middle one-third of college aptitude rating. In other words, if these data in Table IV should hold for students in general, a doubtful assumption because of the small number of the former, counselors do slightly better work with students low in academic aptitude with respect to the improvement of scholarship.

TABLE IV

Relation of Level of Ability and High School Scholarship to Improvement in Scholarship

		1–33 PERCENTILE	34–66 PERCENTILE	67– PERCENTILE
			Advised	
College Aptitude	No.	21	40	11
	Mean	.928	1.14	1.39
			Not Advised	
	No.	22	32	19
	Mean	.618	1.27	1.26
			Advised	
High School Scholarship	No.	5	18	50
	Mean	.26	.853	1.29
			Not Advised	
	No.	5	17	50
	Mean	1.1	.40	1.27
			Advised	
College Aptitude Rating	No.	4	40	28
	Mean	.2	1.08	1.3
			Not Advised	
	No.	3	40	30
	Mean	.7	.84	1.42

Now let us summarize the results of our investigations. We have been attempting to find some satisfactory criterion of the effectiveness of advising, in other words to discover in what ways counseling was helpful to students. We chose as the problem for investigation the improvement in scholarship resulting from such counseling. We found that advisees did no better in scholarship than those students who were not advised. In other words, if counseling is helpful to students, it is not in respect to scholarship that it proves beneficial. We shall have to look elsewhere for criteria of effectiveness, or to other counseling methods than those involved in this study. It seems evident that study motivation is too technical for faculty counselors and should be dealt with by specialized teachers in a How to Study course.

REFERENCES

1. Paterson, D. G. *et al.* "The Minnesota Student Personnel Program." *Educational Record*, Supplement 1928, No. 7, April, p. 35.
2. Paterson, D. G. and Langlie, T. Report of a controlled experiment on the value of faculty advisers for probation students in the College of Engineering, Chemistry, and Architecture. University of Minnesota, 1925–26. Unpublished report.
3. Johnston, J. B. "Predicting Success at the Time of Entrance to College." *School and Society*, 1926, XXIII, January 16.
4. Paterson, D. G. and Langlie, T. "The Influence of Sex on Scholarship Ratings." *Educational Administration and Supervision*, 1926, XII, No. 7, pp. 465–469.
5. Jones, Lonzo. "Personal Service and Freshman Scholarship." *Educational Record*, XII, 1931, January, pp. 71–83.
6. Lemon, A. C. "An Experimental Study of Guidance and Placement of Freshmen in the Lowest Decile of the Iowa Qualifying Examination, 1925." University of Iowa Studies in Education, Vol. III, No. 8, Iowa City, University of Iowa, 1927.
7. Holaday, P. W. "The Long-Time Effect of Freshman Counseling." *School and Society*, XXIX, 1929, February 16, pp. 234–236.

THE EFFECTIVENESS OF APPEAL IN RADIO ADVERTISING: A TECHNIQUE WITH SOME TYPICAL RESULTS

HAROLD V. GASKILL AND RICHARD L. HOLCOMB
Iowa State College

PROBLEM

THIS study is primarily concerned with the development of a technique for evaluating the effectiveness of appeal in radio advertising. We have tried to develop our problem in a way that our results may illustrate possibilities rather than drawing final conclusions. The importance of the problem became apparent when we were initiating several experimental studies of radio advertising. Obviously many aspects of radio advertising are different from similar aspects of printed advertising, so that a technique for experimental attack becomes the subject of investigation. This is a statement of our problem as well as a justification for our paper.

Some of the differences in the variables in a study of radio advertising as contrasted with printed advertising involve fundamental responses. Attracting and holding the attention to radio programs bears a relationship to this same function in printed advertising, but is strikingly different in two respects, namely the memory value for the sponsor or his commodity and the memory value of the program per se. Turning a dial may be more complex or paradoxically more simple than turning a page. In the words of one of our philosophizing radio stars, "Anyone can turn the dial, and I have no use for a person who is too lazy to tune out a bad program." A perceptual response is necessary if a magazine page is turned because of unfavorable or unpleasant advertising, and more meaningful response in terms of memory value is made than in listening to a few seconds of a "bad" radio program which is immediately rejected by turning the dial. When the dial is turned in the situation just described, it is possible that there is no reaction to the commodity being advertised or to the sponsor.

Another aspect of our problem is memory for program content versus advertising script. This again is a distinctly different problem in radio advertising. Also we must ascertain whether or not the individuals have listened to the program—and our problem can be viewed as the same problem one has in a trade test, in which information items are used to determine whether the examinee has been exposed to the trade.

Then, there is the analysis of the different variables and the need for adjustments and corrections in order that comparisons may be made. Incidental to such an analysis are questions of comparing effectiveness of programs for men and women, for age groups, and for occupational groups. The question of the value of individual analyses of errors that are made on program content or on advertising content also arises. And finally, many incidental problems are drawn in by the nature of a study like the investigation being reported.

MATERIAL

Five nationally broadcast programs were selected for investigation. It was assumed that adequate coverage was insured by the large number of chain broadcasting stations in the districts in which persons to be tested reside. All programs were broadcast from relatively powerful stations; and radio reception in Iowa for these particular programs was assumed to be relatively good because of lack of interference for evening programs. All of the programs were broadcast in the evening.

Programs varied in: (1) content and type of entertainment, (2) types of commodities advertised—two of the five programs advertised commodities which sell to men only, (3) time spent on the air, (4) minutes per hour spent in advertising, including introductions in which the commodity or firm name is mentioned, and (5) amount of newspaper and magazine advertising supplementing the broadcasting.

The experimenters listened to the broadcasts of each program being studied for a period of three months to determine the content. No sponsor or station was asked for continuity or

for a description of program content. All programs were sponsored by firms selling one item—like a specific brand of yeast, a specific brand of coffee, etc.,—except one program which was sponsored by a meat packer. This program, however, stressed only two items—hams and bacon, with emphasis on processed hams. Time spent on the air was determined from printed announcements of the programs. The number of minutes per hour spent in advertising was determined by timing announcements with stop watches. The amount of newspaper and magazine advertising supplementing programs was approximated by counting the column inches of advertising space each program sponsor had in a state daily paper of exceptionally good coverage, and the number of pages of advertising in seventeen national weekly and monthly magazines. By inspection, it was evident that the newspapers of one month and issues of magazines covering a two-month period would give an adequate sample of the advertising, in addition to radio broadcasting, for our approximation.

A single choice examination on complete program content was constructed. Two forms were used. Each form contained fifty items, with ten items for each of the programs. Items covered both advertising content and "entertainment" content. Items were approximately equal in length, i.e., in number of words contained. Item difficulty was, of course, random, due to many factors. Each examination pamphlet contained one page of blanks to be filled in with general information by the person being tested. In addition to name, age, education, etc., there were five questions about favorite program, network preference, etc., as well as an estimate of time spent in making advertising announcements for each of the five programs being studied. The examinee was further requested to rank the five programs from one to five in order of their value in advertising, in his judgment, for the company sponsoring each. Finally, a question was asked regarding the amount of time, in terms of times per week and number of weeks, which the examinee spent in listening to each of the five programs. Fol-

lowing this page of questions in the test pamphlet were six and a half pages containing the single choice items, with printed directions for marking the items, including a correctly marked sample. All items were of this general structure:

(1) On advertising content:
A fresh delivery of Blank Coffee is made to your grocer every (1) Week, (2) Fortnight, (3) Month, (4) Ten Days, (5) Two weeks. ( )

(2) On program or "entertainment" content:
The theme song of the Blank is (1) Time on My Hands, (2) Come to Me, (3) Yours and Mine, (4) My Time is Your Time, (5) Now's the Time to Fall in Love. ( )

As noted above, the matter of item difficulty was uncontrolled, except as the sampling for each item might effect a control—and this condition proved to be the main uncontrolled variable of the experiment. We do not want to underestimate the importance of this matter of difficulty of the questions. A sample of five questions on the program content is not necessarily comparable in difficulty with the five questions on the advertising nor are the five questions on one program comparable with five on another program. But we are interested in establishing a technique for investigation at present, and not in establishing final conclusions.

METHOD

The tests were given to available groups of residents of Iowa. Group selections were determined by availability and by an attempt to get a representative sampling which would be random. Accordingly, our groups consisted of college classes, high school classes, service clubs, women's clubs, graduate seminars, and groups of high school students attending conventions on the Iowa State College campus. These convention groups provided an easy way of obtaining a geographical distribution. The service clubs and women's clubs of various towns cooperated willingly—often using the test as an enter-

tainment feature because of the novelty of the situation. All of the tests were administered by the writers.

Complete directions were printed on the forms, so that no verbal directions were necessary. No time limit was set—but the test was not strictly work limit. Twenty minutes after the tests were handed out to each group, the examiner urged all to finish as soon as possible. In no instance were more than thirty minutes allowed. For groups of adults unaccustomed to general test situations, this time limit meant discarding unfinished test pamphlets. Approximately twenty per cent of the pamphlets were discarded because of omissions. Perhaps most of the omissions were due to complete unfamiliarity with the programs. This could be checked by reference to the first page, on which the examinee gave the estimate of amount of time spent in listening to the programs in question. In a few of the tests discarded because of omissions, omissions were due to reticence or refusal of the examinee to attempt to answer the questions. This situation occurred only with the adult non-student groups, and occurred least frequently. For the purposes of this paper, only the data from the completed test pamphlets are used. The criterion for rejection was the omission of one section of any five questions, covering either of the two aspects of any program; viz., entertainment or advertising. Accordingly, data presented here are from 350 of the 485 tests given.

Each item appeared once in each form, and the order of items was reversed in the second form. Table entries in most of our data represent the combined result for the particular question from both forms.[1] After preliminary tryouts the forms were revised and an attempt was made to get comparable structure in each question so that suggestibility would be equated ap-

[1] Question 22 in form I is: The announcer on the Blank hour is 1. Graham McNamee. 2. James Wallington. 3. George Rubinoff. 4. Conny O'Hearn. 5. George Wallington, while the corresponding item appears in form II as question 27: The announcer for the Blank hour is 1. James Wallington. 2. George Rubinoff. 3. Harold McNamee. 4. Joseph Moore. 5. Connie O'Hearn.

proximately (one "Harold McNamee" with one "George Wallington," for example.)

RESULTS

We have chosen a purely arbitrary method of evaluating our results. The analysis of answers to specific questions by means of percentage of correct answers seemed the simplest accurate device for making comparisons. This method of comparison of

TABLE I

Table Entries are Per Cent Correct Answers
"Program" Refers to the Entertainment Portion of the Total Presentation, "Adv." to the Advertising and Selling Portion.
Number of Persons Tested = 350

	MEN					WOMEN				
PROGRAM	Mn College	Mn Psych. & Adv.	Mn Grad. Stud.	Mn Business	Mn Men (N = 174)	Mn Women (N = 176)	Mn College	Mn Psych. & Adv.	Mn Housewives	Mn High School
A	29	24	46	23	*30*	*46*	45	24	60	56
B	58	63	62	42	*56*	*56*	60	58	60	47
C	63	58	66	47	*59*	*43*	36	40	53	42
D	45	32	52	28	*39*	*37*	31	27	51	41
E	51	38	39	33	*40*	*22*	24	18	32	15
ADV.										
A	35	47	52	38	*43*	*39*	37	44	38	38
B	60	69	49	38	*54*	*54*	55	57	59	45
C	35	26	37	39	*34*	*21*	25	9	35	16
D	35	32	39	22	*31*	*21*	19	15	26	25
E	48	35	30	30	*36*	*18*	19	15	22	17

answers to specific questions has been of value in other studies.[2] Our "revisions" (Tables III, IV, V) involve the arbitrary

[2] Burtt, H. E., and Gaskill, H. V., Suggestibility and the Form of the Question, this JOURNAL, 16, 4, 358–373; and, Link, H. C., A New Method for Testing Advertising and a Psychological Sales Barometer, this JOURNAL, 18, 1, 1–26.

manipulations. It is to be remembered that these are presented merely as part of a suggested technique with typical results.

In the discussion that follows, "program" refers to the entertainment or non-advertising portion of the total presentation, and "advertising" to the selling and advertising portion of the presentation. As previously noted, the time of the advertising portion was determined by the writers, who listened to the presentation and timed it with a stop watch. "Presentation" includes both program and advertising.

In Table I, the percentage of correct answers for the sample of 350 population is given as the table entry. Each entry is the averaged score from the two forms of the test, and the mean score for the classification of the persons tested. Answers from men are grouped in the left half of the table, and those from women are in the right half. The upper half refers to questions on the program (non-advertising), and the lower half to questions on the advertising portion of the presentation. A descriptive term in the column heading classifies persons answering as college students, students in courses in Psychology and Advertising, business men, etc. Means obtained by averaging answers to both forms of the test are given for each classification. Each table entry represents an average percentage of correct answers to question,—five from each test form, since each presentation studied was represented by five questions on the program and five on the advertising. Thus the table entry for Program A, mean of college men, is the average of 26 (the per-

TABLE II

Table Entries are Numbers of Persons Tested, by Sex and by Examination Form

Roman Numerals Refer to Form Number

COLLEGE		PSYCH. & ADV.		GRAD. STUD.	BUSI-NESS	HOUSE-WIVES	HIGH SCHOOL	TOTAL	
M	W	M	W	M (only)	M (only)	W (only)	W (only)	M	W
I = 30,	22	I = 23,	12	I = 11	I = 26	I = 20	I = 32	I = 90,	86
II = 22,	19	II = 20,	8	II = 14	II = 28	II = 26	II = 37	II = 84,	90

centage of correct answers to the five questions on Program A by college men, Form I of the test) and 32 (the percentage of correct answers to the five questions on Program A by college men, Form II). N for Form I answers, college men, Program A is 30. N for Form II answers, college men, Program A is 22. (Analysis of N appears in Table II.)

The odd versus even reliability for Form I is .89, and for Form II is .70.

In Tables III, IV, and V, various comparisons are made. Actual performances on parts of the tests are entered, and comparisons between programs are made in terms of ranking

TABLE III
Transmuted Scores

PRESENTATION	PART A				PART B		
	Program Av. Score	Program	Adv. Av. Score	Adv.	Program and Adv.	Adv.	Program
A ____	40	75.5	39	81.2	78.3	81.2 M & W	75.5
B ____	53	100.0	48	100.0	100.0	100.0 M & W	100.0
C ____	50	94.3	29	60.4	77.3	60.4 M & W	94.3
D ____	36	67.9	26	54.2	61.0	64.8 M	67.9
E ____	30	56.5	27	56.2	56.4	75.0 M	56.5

Program refers to the non-advertising portion of the total presentation; adv. to the advertising and selling portions of the presentation.

by a percentage transmutation. Thus, in Part A of Table III, the raw score means from the part of the tests covering program content are given in the column *Program Av. Score;* and the entries in the *Program* column are the transmuted scores, obtained by setting the highest ranking score at 100 (in this instance the score for presentation B) and multiplying each of the others by 100/53. Entries in the column headed *Program and Adv.* (Part A, Table III) are merely the sum of the entry in *Program* and the entry in *Adv.*, divided by two. This gives a transmuted score for the percentages of correct answers on the program and advertising combined.

Commodities advertised in presentations D and E sell almost exclusively to men (cigars and smoking tobacco). Using

only men's replies for D and E, but replies from both men and women for A, B, and C, Part B of Table III is developed. However, selection of data on the basis of reply by sex is pertinent only to column 1 of Part B, *Adv.*, since program content must be appealing to both men and women to insure listeners; therefore scores of both men and women are used in column 2, *Program*.

The programs vary in the amount of time spent on the air. Solely to obtain a comparison, we assume a one-to-one relationship between effectiveness and time spent on the air. With one quarter hour per week as a unit, adjustment is made for the scores of Table III, and these adjustments and revisions appear in Table IV. Thus, scores in the *Program* column are divided by the number of quarter hours per week of broadcasting for that program; then the highest ranking is set at 100 (Presentation D, which has average score of 34.0 for program adjusted to number of quarter hours) and the others are given a percentage relationship. These percentage revisions appear as columns headed *Program Revised* and *Adv. Revised* in Table IV.

Again an assumption is made solely to obtain a comparison—the assumption that the effectiveness of advertising is directly proportional to the time spent in advertising, e.g., the time during the total presentation spent in advertising. We correct for this factor by dividing the *Adv. Revised* scores of Table IV by the time spent in advertising. Then the highest ranking is set at 100 and the others adjusted in percentage relationship. This appears in the last three columns of Table IV. Only one presentation (E) uses a definite "build-up," designed to develop a favorable and receptive attitude toward the presentation. Accordingly, E is revised both to include and to omit this "build-up"—hence two entries for E.

Seventeen magazines and a widely circulated daily newspaper were scanned for advertising by the sponsors of the radio presentations under consideration. The data on the amounts of this printed advertising appear in Table VI. No

TABLE IV
Revisions; Time Spent on the Air

PRESENTATION	PROGRAM	PROGRAM REVISED	ADV.	ADV. REVISED	MIN. PER WEEK PRESENTATION (INCL. PROG. & ADV.)	SEC. ADV. PER PRES.	TIME ADV. PER HOUR	REVISED TIME	ADJUSTED REVISION
A	15.1	44.4	16.2	50.0	75	120	8.0	5.5	19.1
B	25.0	73.6	25.0	77.1	60	190	3.16	23.0	79.3
C	23.6	69.1	15.1	47.7	60	195	3.25	21.0	72.4
D	34.0	100.0	32.4	100.0	30	105	3.5	29.0	100.0
E	9.4	27.6	12.5	38.6	72	180 or 65	12.0 or 4.33	2.3 or 6.4	7.9 or 22.1

adjustment is made for page position, circulation of the medium, or other similar factors. However, commodity A is the only commodity represented by printed advertising in four colors, with all ads in exceptionally good locations within the medium, so that the value of the printed advertising for A is adjusted arbitrarily. It is to be remembered that all broadcasts were evening broadcasts—all quite near 7:00 P. M., C. S. T.

For evaluating our technique, assuming the questions to be of equal difficulty, several tentative conclusions may be drawn. In general, the presentations are successful in getting facts about the product and sponsor to the audience. In most presentations, however, the memory value of program content is higher than the memory value of the advertising content. Certainly there appears no relationship between actual time devoted to advertising and its memory value (Table IV). Even though a presentation is designed to sell a commodity to men only, women are comparatively well informed about the presentation. (Table III). Presentation B ranks first in memory value for both men and women,—both in advertising content and in program content. Ranking second in advertising alone (both men and women) is A, but second in program is C. Both A and C are general commodities—A, hams and bacons, and C, coffee. The fifth ranking, C, for advertising, as noted above, is second in program content—and in program content fifth position is taken by E, a smoking tobacco.

More detailed comparisons are possible when one refers to Table I. Data from answers by men appear in the left half, and from answers by women in the right half. The two middle columns give the means.

When Tables I and III, IV, and V are compared with Table VI, it is evident that the preponderance of printed advertising for commodity B bears a relation to B's position in Table III. When data for D and E, neither of which has printed advertising, are adjusted (Table V) the rankings change.

TABLE V

Revisions, Including Printed Advertising

PRESENTATION	PRINTED ADV.	REVISED	BROADCAST AND PRINTED ADV.	PROGRAM	ADV. AND PROGRAM
A	6	3.2	3.2	44.4	23.8
B	20	3.9	3.9	73.6	38.7
C	8	9.1	9.1	69.1	39.1
D	0	100.0	100.0	100.0	100.0
E	0	7.9 or 22.1	7.9 or 22.1	27.6	17.7 or 24.8

Our specific method of correction, being arbitrary, may not be appropriate in each instance, but it does give a comparative basis for effectiveness of appeal and for memory value. We can use units in which the broadcaster is interested—percentage of advertising per minute on the air.

May we again emphasize the difficulty of the questions in the test pamphlets as the main uncontrolled variable. This is true both for advertising content and for program content. We have suggested a control for this by using a fair sized sample of questions. No attempt was made to isolate the amount of carry-over from older published advertising, or to ascertain the exact amount of influence of current printed advertising. Neither of these points is pertinent to the main problem being investigated.

INCIDENTAL PROBLEMS

This technique might well be extended to problems such as the relative merit of musical versus dramatic programs. Comparisons could be made between broadcasts at different times of the day, or of the week.

A further study of the exceptionally high memory value of ingeniously "woven in" advertising announcements might afford new insight into the problem of remembering. Another possibility would be to attempt to equate by having the subject

TABLE VI

	A	B	C	
S. E. P.	3 Full pg. color	4 Full pg. BW	2 Full pg. BW	
Colliers		2 Full pg. BW ½ pg. BW ¼ pg. BW		
Ladies Home Journal	1 Full pg. color	1 Full pg. BW	2 Full pg. BW	
Good Housekeeping	1 Full pg. color	2 Full pg. BW ⅛ pg. BW	2 Full pg. BW	
Literary Digest		2 Full pg. BW 3–⅔ pg. BW 1–⅓ pg. BW		(D and E, none)
Physical Culture		2 Full pg. BW		
Des Moines Register		264 Col. Inches As: 4–21 in. 3–60 in.	100 Col. Inches As: 1–40 in. 1–60 in.	

Journals were for April and May. In addition to those noted, issues of the American, Field and Stream, Popular Science, Popular Mechanics, Wallace's Farmer, Bureau Farmer, Photoplay, Radio News, Radio Digest, Farmer's Wife, and Country Gentleman, were inspected. No advertising for any of the commodities appeared in these.

listen to a specific broadcast on some other pretext and then take the test. Question difficulty could then be determined on the basis of a comparison of percentage of correctly answered questions.

The comparative frequency with which different selling points have gone over to individuals, all of whom have listened to the broadcast, could be analyzed. We could then determine whether such points have been remembered because of interest or because of repetition, or place within the program.

SUMMARY

Tests, made up of fifty single choice items on complete program content of five nationally broadcast programs, were given to groups of business men, college students, women's clubs, and high school students. N = 350.

Programs varied in: (1) time spent on the air, (2) minutes per hour spent in advertising, (3) types of commodities advertised, and (4) the amount of newspaper and magazine advertising in addition to the broadcasting.

Typical results, which must be viewed as tentative conclusions, in the light of our discussion above, are:

1. More program content than advertising content is remembered. Points like "qualities that ovenizing gives to hams and bacons" are missed by more than eighty per cent of the persons tested. There is no relation between amount of time spent in advertising and its memory value.
2. Men were better informed than women about the programs of the two commodities advertised directly to men, but were likewise better informed than women about a program advertising a commodity to both sexes.
3. Greater memory value for advertising content seems to be directly affected by the ingenious "sandwiching in" of advertising announcements. Complete isolation of advertising announcements, such as introductions and

conclusions distinctly separated from the program seem to make for low memory value.

4. Newspaper and magazine advertising has a negligible influence—but our five programs give an entirely inadequate basis for generalizing this conclusion.

Various extensions of this technique could produce interesting results. The time of day, question of dramatic and musical programs, the number and power of re-broadcasting outlets, and similar problems are amenable to treatment by this technique.

PERSONALITY AS RELATED TO BIRTH ORDER AND FAMILY SIZE

ROSS STAGNER
University of Akron

AND

E. T. KATZOFF
Chicago, Illinois

PSYCHOLOGICAL literature is full of statements as to the effect of being an only child, an oldest child, a youngest child, etc., upon the personality. Unfortunately, the evidence upon which these statements are based is conspicuous mainly by its absence. The fallacies of personal impressions, selected cases and unwarranted interpretations are nowhere in psychology so well illustrated as here.

Stuart (6), using the Laird B 2 and C 2 inventories, reports that family ranking did not show a significant relationship to personality test score. Unfortunately, he gives no indication of the number of cases, the size or direction of the differences, etc. The tests used are low in reliability and of questionable validity.

Bender (1) employed the Allport A-S Reaction Scale with a group of college men. His observations indicated that only children are likely to be more ascendant and youngest children more submissive than the average.

Various studies of delinquency, such as that of Wile and Noetzel (7), infer from the occurrence of certain types of behavior problems in association with certain family constellations (only child, second child, etc.) that this proves an effect on personality. Their results, however, might be interpreted as showing the same personality reacting to different social environments. Other studies on delinquents do not agree with Wile and Noetzel that the frequencies of problems are related to birth order or family size.

Maller (3) reports that size of family is not related to such factors as moral knowledge or honesty as measured by the C.E.I. tests. Average-sized families gave best scores on co-operativeness, inhibition and helpfulness. Large families scored best on persistence. The general results show children of large families at a disadvantage, but the differences were small.

In view of the confused nature of these observations, no apology is needed for another study using a well-standardized personality scale and a larger number of cases than the studies listed.

METHOD

The Bernreuter Personality Inventory was administered to 430 men attending the University of Wisconsin in the fall semester 1932–33. About half were freshmen. All indicated the number of children in their families, and the subject's ranking in this number.

The low validity of scale B3-I prompted us to omit it and use only scales N (emotionality), S (self-sufficiency) and D (dominance). Data on birth-order and family size are reported for these three scales.

Six groups were formed: only children; older of two; younger of two; oldest of three or more; youngest; and intermediate. Re-grouping, the first three groups made up a "small family" group, contrasted with the last three, "large family" cases.

To study the effect upon the personality of the child of being "dispossessed" by having another child born into the family after him, another re-grouping included only children, younger of two, and youngest of several (last-born of their families), contrasted with older, oldest and intermediates, all of whom were supplanted in the "baby" position by a later-born child.

RESULTS

Table I gives the mean scores and sigmas for all of the six original groups. Differences and critical ratios on scale N

(Table II), scale S (Table III) and scale D (Table IV) summarize the findings on birth order. Direction of the difference in II, III or IV can be determined by referring to Table I.

TABLE I
Means and Sigmas on all Groups[1]

GROUP	MEAN			SIGMA			CASES
	N	S	D	N	S	D	
Only	801.1	919.4	927.4	71.7	41.2	62.4	57
Older	791.0	911.1	930.8	77.1	51.6	62.4	67
Younger	802.4	898.3	925.9	74.3	43.5	52.1	42
Oldest	819.5	901.5	916.7	82.8	45.1	55.6	72
Youngest	831.1	895.4	911.4	77.3	47.1	59.1	73
Intermediate	817.6	903.7	914.4	81.7	52.8	63.9	119

[1] These mean scores are 875 points higher than Bernreuter's, because of our time-saving scoring method (4).

Table II shows that the older of two (most stable) and youngest of several (most unstable) give the only critical ratios above 2.00 on scale N. Only one C.R., that comparing these two directly, reaches the arbitrary criterion of 3.00.

TABLE II
Differences and Critical Ratios on Scale N
(Upper figure is the difference between the two means; lower figure, the critical ratio—difference/standard error.)

	OLDER	YOUNGER	OLDEST	YOUNGEST	INTERMED.
Only	10.1 .75	1.3 .08	18.4 1.35	30.0 2.28	16.5 1.36
Older		11.4 .76	28.5 2.09	40.1 3.06	26.6 2.60
Younger			17.1 1.13	28.7 1.91	15.2 1.11
Oldest				11.6 .86	1.9 .15
Youngest					13.5 1.14

Table III shows a similar situation with reference to self-sufficiency. The only children are most self-sufficient, the youngest of several are lowest. The C.R. for these two is

TABLE III

Differences and Critical Ratios on Scale S

(Upper figure is the difference between the two means; lower figure, the critical ratio—difference/standard error.)

	OLDER	YOUNGER	OLDEST	YOUNGEST	INTERMED.
Only	8.3	21.1	17.9	24.0	15.7
	.99	2.34	2.34	3.08	2.07
Older		12.8	9.6	15.7	7.4
		1.39	1.16	1.87	.90
Younger			3.2	2.9	5.4
			.37	.34	.63
Oldest				6.1	2.2
				.80	.30
Youngest					8.3
					1.09

above 3.00, and the only child is high enough on self-sufficiency to score C.R.'s of 2.00 or more when compared with younger of two, oldest and intermediate of several. The value of being a first-born child, as far as this trait is concerned, does not extend to the oldest of several, this group ranking fourth of the six groups.

Table IV shows that no groups even give a critical ratio of 2.00 on dominance. This may be due in part to the lower validity of scale D.[2]

Table V presents the mean scores and differences when the cases were divided into only two groups, the small family (one or two children) and the large (three or more). The advantage of the small family child as to emotionality and dominance is apparent although not perfectly significant.

Table VI compares those children who have been "dispossessed" (earlier-born) with those who have not (later-

[2] As judged by interviews with the subject (5).

TABLE IV

Differences and Critical Ratios on Scale D

(Upper figure is the difference between the two means; lower figure, the critical ratio—difference/standard error.)

	OLDER	YOUNGER	OLDEST	YOUNGEST	INTERMED.
Only	3.4	1.5	10.7	16.0	13.0
	.30	.13	1.00	1.48	1.25
Older		4.9	14.1	19.4	16.4
		.44	1.40	1.89	1.66
Younger			9.2	14.5	11.5
			.89	1.37	1.13
Oldest				5.3	2.3
				.55	.25
Youngest					3.0
					.32

TABLE V

Comparison of Small and Large Families

GROUP	MEAN			SIGMA			CASES
	N	S	D	N	S	D	
Small	799.0	910.7	928.4	76.8	74.2	65.8	166
Large	821.9	900.8	914.2	82.6	62.3	63.1	264
Diff.	22.9	9.9	14.2				
C. R.	2.91	1.42	2.20				

TABLE VI

Comparison of Earlier-Born and Later-Born Children

GROUP	MEAN			SIGMA			CASES
	N	S	D	N	S	D	
Earlier-born	803.2	913.9	925.2	77.7	48.2	66.9	234
Later-born	819.1	900.5	915.5	95.3	52.8	68.1	196
Difference	15.9	13.4	9.7				
C. R.	1.91	2.74	1.48				

born). The advantage of *not* being dispossessed (since the literature usually refers to the injury done the child's per-

sonality when his supremacy is usurped) disappears in this table. In fact, one might infer that continuing in the "baby" position leads to a lack of independence which is rather undesirable.

DISCUSSION

The fact that the findings of the present study are largely negative should not be surprising to anyone who has worked with personality problems. The surprising point is that so many presumably thoughtful psychologists have emphasized the importance of birth order in determining personality.

It is obvious that we can rationalize the findings of any differences in personality in terms of family constellations. If the only child proves too nervous, we say this is due to his lack of normal social contacts. If he proves less nervous than the average, we say it is because he has had fewer jealousies and emotional traumata than the child with brothers and sisters. If he proves aggressive and independent, we explain it on the basis of his supremacy in the home, and if he is submissive and dependent, that is because he leaned too much on his parents.

The same process of analysis could be applied to the statistical findings of our investigation. It would not enhance the knowledge of personality development in any way. The determinants of personality traits are complex interactions of heredity and environment. Into this interaction birth order enters only as an insignificantly small factor, which may (as Bender noted) appear statistically "significant" with this population and insignificant with another.

The desirability of the small family has more logic on its side, but this logic is in the direction of superior economic, educational and social advantages, not in the size of the family *per se*.

SUMMARY

1. The results of testing 430 college men with the Bernreuter Personality Inventory indicate that differences in birth

order, with few exceptions, do not result in significant differences in personality score.

2. A slight advantage for small families is indicated.

3. The effect of being dispossessed by a later-born child seems to be the development of more independence.

4. The effect of family constellation should be considered as only a small portion of the continuous interaction which determines personality.

BIBLIOGRAPHY

1. Bender, I. E. Ascendance-submission in relation to other factors in personality. J. Abn. & Soc. Psychol., 1928, *23*, 137–43.
2. Crane, G. W. Psychology Applied. Chicago, 1932. (pp. 445–448).
3. Maller, J. B. Size of family and personality of offspring. J. Soc. Psychol., 1931, *2*, 3–27.
4. Stagner, R. Scoring weighted-items tests: a short-cut method. Personnel J., 1932, *11*, 174–176.
5. Stagner, R. Validity and reliability of the Bernreuter Personality Inventory. J. Abn. & Soc. Psychol., 1934, *28*, 413–418.
6. Stuart, J. C. Data on the alleged psychopathology of the only child. J. Abn. & Soc. Psychol., 1926, *20*, 441.
7. Wile, Ira S., and Noetzel, E. Study of birth order and behavior. J. Social Psychol., 1931, *2*, 52–71.

THE RELATIONSHIP BETWEEN FINGER LENGTH, HAND WIDTH AND MUSICAL ABILITY

JAMES H. TAYLOR
Wittenberg College

POPULAR psychology has long urged that an intimate and causal relationship exists between the formation of the hands and the proficiency of music students. This study was undertaken to evaluate the assertion that musicians possess unusually long fingers and narrow hands, and the claim that musical proficiency varies in proportion to finger length, or hand width, or some combination of the two.

In order to obviate the structural difference between male and female hands, and because there appear to be more women than men music students, only women were used as subjects. Hand measurements were made on 100 students, including 40 pianists, 30 violinists and 30 non-musicians who formed the control group. The first 20 pianists, the first 14 violinists and the first 20 non-musicians were students at The Ohio State University (see Table I). The remaining subjects were recruited from the liberal arts and music schools at Wittenberg College.

The first 20 pianists were selected and ranked[1] from 1 to 20 on the basis of playing proficiency by Dr. Wilson and other members of the music staff at Ohio State. The individual ranked 1 is considered as the best woman pianist then receiving training at the University; the individual ranked 20 is considered as the least able of this select group. Professor Eugene Wiegel, director of the Ohio State University Sym-

[1] While this method of rating ability is admittedly open to error, it is the only feasible and the most practical method available. Insofar as musical ability can be estimated by experts in the field of musical training and supervision, the rankings as set by Professors Wilson and Wiegel are adequate.

TABLE I

Finger Lengths, Hand Widths and L/W Ratio for Pianists (P), Violinists (V) and the Control Group (C)

	FINGER LENGTH IN CENTIMETERS			HAND WIDTH IN CENTIMETERS			FINGER LENGTH/HAND WIDTH (L/W)		
	P	V	C	P	V	C	P	V	C
1	11.1	11.0	11.9	8.0	7.3	7.4	1387.5	1506.8	1472.9
2	11.0	11.7	10.9	7.3	7.2	7.3	1506.8	1625.0	1493.1
3	10.5	10.5	11.2	7.5	7.6	7.8	1400.0	1381.5	1435.8
4	11.1	10.8	9.8	7.8	8.0	6.8	1423.0	1350.0	1441.1
5	10.7	9.5	10.5	7.5	7.8	7.5	1426.6	1269.2	1400.0
6	11.6	11.1	10.3	8.1	7.1	7.1	1432.1	1563.3	1450.7
7	11.8	10.3	10.8	8.1	7.1	7.5	1456.7	1450.7	1440.0
8	11.3	11.2	10.9	7.9	7.0	7.8	1430.3	1600.0	1397.4
9	10.9	10.4	10.3	8.0	7.3	7.3	1362.5	1424.6	1471.4
10	11.2	10.0	11.1	7.5	7.3	7.5	1493.3	1369.8	1480.0
11	10.8	10.6	10.4	7.8	7.6	7.5	1384.5	1395.7	1386.6
12	11.1	9.8	10.0	7.6	6.8	7.0	1460.5	1441.1	1428.5
13	10.8	11.3	10.3	7.3	8.1	7.2	1479.4	1395.0	1430.5
14	10.0	9.9	10.4	7.5	7.1	7.4	1277.1	1394.3	1405.4
15	10.8	9.8	10.0	7.3	7.1	7.2	1333.3	1340.0	1388.8
16	10.6	11.3	10.7	8.3	7.5	7.6	1486.4	1506.7	1407.8
17	12.0	11.3	10.9	9.0	8.3	7.5	1401.2	1361.4	1453.3
18	11.0	10.5	10.3	7.4	7.5	7.0	1325.0	1400.0	1471.4
19	11.4	10.4	10.5	8.2	7.3	7.2	1479.4	1424.7	1458.3
20	10.6	11.0	10.0	8.0	7.5	7.5	1325.0	1466.7	1333.3
21	10.2	10.0	10.9	7.3	7.8	7.4	1397.3	1400.0	1378.4
22	11.5	11.5	11.2	8.0	7.8	7.8	1437.5	1474.3	1435.9
23	10.4	10.4	10.5	7.6	7.5	7.5	1368.4	1368.7	1400.0

TABLE I—(*Continued*)

	FINGER LENGTH IN CENTIMETERS			HAND WIDTH IN CENTIMETERS			FINGER LENGTH/HAND WIDTH (L/W)		
	P	V	C	P	V	C	P	V	C
24	10.6	10.4	9.9	7.4	7.4	7.1	1433.8	1405.4	1394.4
25	10.5	11.4	10.3	7.9	7.7	7.3	1329.1	1480.5	1411.0
26	10.5	11.2	11.1	7.9	8.2	7.9	1329.1	1365.8	1405.1
27	10.4	11.1	10.4	7.5	7.7	7.1	1386.6	1441.6	1464.8
28	9.7	10.9	10.7	8.0	7.5	7.2	1212.5	1405.3	1486.1
29	10.9	10.9	10.3	7.6	7.5	7.0	1434.2	1405.3	1471.4
30	10.1	10.5	10.9	7.1	7.4	7.6	1408.4	1419.1	1434.2
31	10.9	.....	.....	7.9	.....	.....	1380.0	.....	.....
32	10.1	.....	.....	7.4	.....	.....	1364.9	.....	.....
33	11.1	.....	.....	7.8	.....	.....	1423.0	.....	.....
34	10.5	.....	.....	7.4	.....	.....	1419.0	.....	.....
35	10.7	.....	.....	7.5	.....	.....	1426.6	.....	.....
36	11.3	.....	.....	7.6	.....	.....	1486.8	.....	.....
37	11.2	.....	.....	7.8	.....	.....	1435.9	.....	.....
38	10.8	.....	.....	7.4	.....	.....	1459.4	.....	.....
39	10.5	.....	.....	7.3	.....	.....	1439.7	.....	.....
40	10.8	.....	.....	7.5	.....	.....	1440.0	.....	.....
Total	432.9	322.0	316.4	308.0	225.0	220.7	56282.8	42750.4	42927.6
Mean	10.8	10.73	10.54	7.7	7.5	7.36	1407.7	1425.0	1430.9
S.D.	0.485	0.525	0.383	0.394	0.355	0.276	61.4	80.6	51.3
P–C	0.26	V–C	0.19	0.34	.....	0.14	C–P– 23.2	.....	C–V– 5.9
CR	2.1	.....	1.65	4.6	.....	1.7	CR– 1.72	.....	CR– 0.34
	98/100	.....	95/100	999/1000	.....	96/100	96/100	.....	63/100

phony Orchestra, selected and ranked all of the 14 women violinists playing under his direction during the winter quarter of 1934. Since performance ratings on the Wittenberg students were not obtainable, only the OSU group was used in determining correlations between hand formation and musical ability. Both groups were thrown together for all other measurements.

As a control and comparative group 30 women who were not classed as musicians were used. Since it is almost impossible to find a college student who has not received some musical training at some time during her life, it was necessary to limit membership in the control group to those who had received no musical training in the preceding five years, and who, at the time of the measurements, were not engaged in either vocal or instrumental work.

Hand measurements were all made on the left hand, since this is the hand most used by violinists. Fortunately, no member of the violinists' or pianists' group was left-handed, and no left-handed subjects were included in the control group. Finger length was measured in millimeters from the tip of the third finger to the apex of the third finger knuckle, including the three phalanges. This measurement was made while the fingers were bent straight out from the third phalangeal joint, forming a right angle to the hand surface. Due allowance was made for fingernail length by subtracting from the finger length measure the amount the nail extended over the finger tip. Naturally, neither the pianists nor violinists boasted of protruding finger nails. Hand width was measured from the knuckle on the index finger to the knuckle on the little finger. In all cases "free" measurements were made (such as are taken in measuring a rod with calipers) so that factors such as extreme boniness or puffiness of the hands could be accounted for. In order to consider the relationship between finger length and hand width, the ratio finger length divided by hand width (L/W) was used. For convenience in handling the data, this ratio is multiplied by 1000.

THE RESULTS

The results appear in Table I. The mean finger length for the pianists (P) is 10.8 cm., for the control group (C) 10.54. The difference of 0.26 cm. expresses a critical ratio (cr) of 2.1, or 98 out of 100 chances of there being a true difference. The mean hand width for P is 7.7 cm., for C, 7.36. This difference is 0.34, with a cr of 4.6, or practically 100/100 chances. The mean L/W for P is 1407.7, for C, 1430.9. The difference of 23.2 expresses a cr of 1.7 or 96/100 chances.

The mean finger length for the violinists (V) is 10.73 cm., for C it is 10.54. The difference of 0.19 is not significant, the cr being only 1.65 or 95/100 chances. The mean hand width for V is 7.5, for C, 7.36. This difference is 0.14, with a cr of 1.7 or 96/100 chances. The mean L/W for V is 1425.0, for C, 1430.9. The difference of 5.9 expresses a cr of 0.34, or only 65/100 chances.

The rank difference correlations between the pianists' rank order of playing proficiency and finger length, hand width and L/W are $.15 \pm .15$, $-.15 \pm 15$, and $.31 \pm .14$ respectively. The first two are of just chance order, while the third suggests only a remote and an unreliable relationship. The correlations between the ranked order of violin playing ability and finger length, hand width and L/W are $.38 \pm .16$, $.18 \pm .20$, $.18 \pm .20$ respectively. While the first R indicates a slightly positive, though unreliable relationship, the latter two are no more than chance.

CONCLUSIONS

Our data indicate that pianists' hands are significantly wider than the hands of the control group; that there is a fairly significant difference in the length of the fingers and in the ratio found by dividing the finger length by the hand width. However, our results belie the usual assertion that pianists' hands are longer and narrower than normal hands. The fingers are slightly longer, but the hands are also wider, indicating a general, all-over growth, rather than a uni-

directional development. The same general conclusions are applicable to the violinists, save that the differences between the violinists and the controls are not so great.

The low correlations indicate little or no relationship between hand formation and estimated proficiency. While the number of cases is small, the rank-difference method weak, and the criterion of ability rather untrustworthy, our data would suggest the inadvisability of predicting success in music on the basis of hand conformation.

When it is considered that the difference between the pianists' hand width and the control hand width is only 34 millimeters, and that this difference is the only statistically significant difference of them all, then it is apparent that only fine measurements will disclose hand growths which have, historically, been reported after casual observation. It would be the keen-sighted individual who could pick out pianists on the basis of hand formation!

Since the differences are so slight, it is debatable whether they may be ascribed to chance, hereditary or training factors. It is quite possible that any differences which do appear are due to the daily hand exercise and practice which are so essential in the manipulation of the piano and the violin. The fact that the hands of both pianists and violinists show an all-over growth would support this suggestion. We have some data, insufficient to tabulate, to suggest that the violinists' left, and most-used hand, is slightly larger than the right, or bow, hand.

REACTION TIME IN AUTOMOBILE DRIVING

BRUCE D. GREENSHIELDS
Denison University

AN automobile does not become a complete mechanism until the driver has placed himself behind the wheel. It is the driver who sees the curve ahead and turns the wheel accordingly, who senses the roughness of the road and lessens the pressure of his foot on the accelerator, and who hears the policeman's whistle and brings the brakes into action. The quickness with which the driver can react to any situation becomes a very important factor in driving especially at high speeds. The brief interval between seeing, hearing, or feeling and the starting to do something in response to the stimulus is called "reaction time." In order to point out to the public its significance, the Safety Division of the Traffic Bureau of the Ohio State Highway Department designed and constructed apparatus for measuring brake reaction-time.

This apparatus was exhibited at the Ohio State Fair, August 1935 and later, in the fall, at the Columbus Auto Show. Some interesting information was brought to light. Brake reaction-time varies with the distance the foot has to travel from the accelerator to the brake pedal. Race track drivers and others who use the left foot to actuate the brake, and travel with it in contact with the brake pedal can thereby react more quickly. A record of driving experience can be obtained by means of a questionnaire.

Mr. R. Z. Moore, Safety Engineer of the Traffic Bureau, not only supervised the project but helped to design the apparatus and conduct the test. Dr. S. E. Haven, assistant Professor of Psychology, Ohio State University, helped to design the questionnaire and made suggestions as to the method of analysis.

The apparatus consisted essentially of the facsimile of the front of an automobile, a short miniature roadway containing an intersection with traffic light, a toy automobile which ran on the road, and a chronoscope (or stop watch) reading to 1/100 of a second. (See Figure 1)

In taking the test to determine his reaction time the subject placed himself in the seat facing the miniature traffic light on the roadway. He was instructed to depress the accelerator whereupon the green light in the traffic signal came on and a bell started to ring. This assured the observer that the person taking the test had the accelerator down and that he was ready to react to the flashing of the red light in the traffic signal.

The test supervisor, after the subject had his foot on the accelerator for two seconds or longer to allow him to get ready, pressed a button which simultaneously changed the traffic light from green to red, started the stop watch and released the toy car at the top of the small incline so that it started to roll down the roadway toward the traffic signal. The person taking the test when the red light flashed removed as quickly as possible his foot from the accelerator and depressed the brake which broke electrical circuits to stop the chronoscope, and turn off the traffic light. A center strip in the road raised and stopped the toy car. The place where the miniature car stopped showed the approximate reaction time. The reaction time was then read and recorded. When the car was replaced at the top of the incline, all was in readiness for another cycle of operations for a new reading. The car helped to direct attention to the test. At the Auto Show, a large dial reading to $\frac{1}{5}$ of a second was added.

The exhibition attracted considerable interest and several thousand people took the test. Over 1400 people filled out the questionnaire giving their ages, driving experience, occupation and other information. The reaction time of these persons was recorded and analyzed. The age at which these

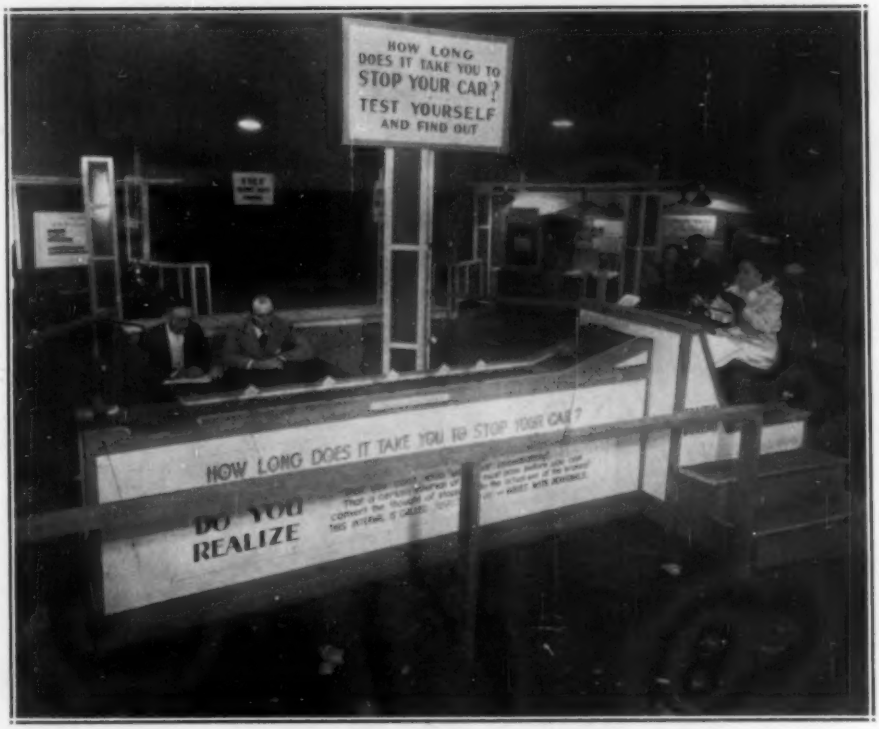

HOW LONG
DOES IT TAKE YOU TO
STOP YOUR CAR?
TEST YOURSELF
AND FIND OUT
HOW LONG DOES IT TAKE YOU TO STOP YOUR CAR?
DO YOU
REALIZE

FIG. 1.

persons started driving and the number of years they had driven are given in the following table.

DRIVING EXPERIENCE

AGE STARTED TO DRIVE	NUMBER	NUMBER OF YEARS' DRIVING EXPERIENCE	
Under 14 years	601	Less than 1 year	28
15–18 years inclusive	572	1– 5 years	536
19–22 years "	96	6–10 years	374
23–25 years "	36	11–15 years	238
26–30 years "	56	16–20 years	193
31–35 years "	34	All over 20 years	65
36–40 years "	19	No data	27
41–50 years "	20		
51–60 years "	6	Total	1461
Over 60 years	1		
No data	20		
Total	1461		

The average brake reaction time for the 1461 persons who took the test was .496 seconds; the standard error of the mean:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}} = \frac{.0913}{\sqrt{1461}} = .00241 \text{ seconds.}$$

In general 5 readings were taken and the average of the mid-scores was .492 seconds with a standard error of the mean, $\sigma_{\bar{x}} = \frac{.0884}{\sqrt{1461}} = .0023$ seconds. This indicates that perhaps the middle value of the 5 readings gave just as good a score as the average of all five.

Certain facts about reaction time have already been established and these need only to be mentioned in this paper. 1. The speed of all forms of reaction vary from one person to another and from time to time in the same person. 2. Reaction time changes gradually with age, very young and very old people being slower in their reactions. The average simple reaction time for women is a trifle less than for men. 3. People generally react more quickly to very strong stimuli than to weak ones. Complicated situations take longer to react to than simple ones. 4. A person's physical condition affects his reactions, fatigue for instance, lengthens them.

5. Distractions increase the time of all reactions except the reflex.

In a previous test conducted at Michigan University in 1933, the writer found that an inconvenient location of the brake pedal in reference to the accelerator caused an average difference of .2 second for four individuals on two different cars.[1] The average time of 13 persons tested on one car was 0.86 seconds, and the average of 27 persons tested on the other car was 0.74 seconds. In two different relative locations of brake and accelerator in the present investigation it was found that one resulted in a brake reaction time of .529 seconds and a standard error of the mean

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}} = \frac{.1017}{\sqrt{250}} = .00643 \text{ seconds.}$$

The average reaction time for the other location with 181 tests was .458 seconds. The standard error of the mean

$$\sigma_{\overline{X}} = \frac{.1224}{\sqrt{181}} = = .0091 \text{ seconds.}$$

In the first position the movement of the ball of the foot from the depressed position of the accelerator was 5″ up and 7¼″ over and in the second position 5″ up and 1″ over. The significant differences of the two results was

$$\sigma_d = \sqrt{\frac{\sigma 1^2}{N_1} \times \frac{\sigma 2^2}{N_2}} = .01114 \text{ seconds.}$$

The actual difference of .071 seconds was more than three times this proving that the difference was a real one.

Another interesting bit of information was furnished by 9 drivers who stated that in their driving of ambulances or racing cars they had learned to drive with the left foot on the brake in order to have better control of the car. The average reaction time for these drivers was .294 seconds, and

[1] The Photographic Method of Studying Traffic Behavior, *Proceedings of the Thirteenth Annual Meeting of the Highway Research Board 1933*, pp. 383–96, by Bruce D. Greenshields.

$$\sigma_{\overline{X}} = \frac{.0559}{\sqrt{9}} = .0182 \text{ seconds.}$$

The longest reaction time was .436 seconds and the shortest .233 seconds. One of these drivers had a reaction time of .65 seconds in the normal way of driving as against .30 seconds, and another .47 and .28 seconds, respectively.

These results naturally raise the question as to whether manufacturers have done everything possible to improve braking control. The hand is quicker than the foot. Why should not power brakes be operated by finger control?

For the Auto Show a more attractive type of questionnaire was used. The questions were all contained on the inside of a folder which the test operator did not open in the presence of the person taking the test, but wrote the results of the test on the back. Of 284 questionnaires filled out at the Auto Show 13 drivers admitted having been in accidents involving no personal injury; 70 in ones involving personal injury, and 4 in accidents involving fatal injuries. One hundred forty-five reported no accidents over a period of five years and 52 furnished no information. The fact that approximately $\frac{1}{3}$ admitted having been involved in accidents indicates that it is possible to secure records of driving experience. With these records at hand it should be possible to develop a criterion for driving ability.

The tests for which the accident records were furnished were too few in number to give definite proof, but they indicate that slow reaction time may have little to do with accident proneness. The 87 persons with accident records had an average reaction time of .4775 seconds, and

$$\sigma_{\overline{X}} = \frac{.079}{\sqrt{87}} = .0085 \text{ seconds}$$

and the 145 with no accidents had an average reaction time of .4768 seconds, and

$$\sigma_{\overline{X}} = \frac{.096}{\sqrt{145}} = .0079 \text{ seconds.}$$

To conclude, the most important things brought by the investigation were that the public is interested in automobile driving to the extent that it will lend willing coöperation if properly appealed to; that a convenient location of brake and accelerator pedal is important; and that, in general, drivers can compensate for slow reaction time by driving more carefully. It may well be that driving tests are of more value in training drivers than in revealing their inability. Of paramount importance is the checking of driving tests against driving experience.

THE MEASUREMENT OF ATTITUDES TOWARD PARENTAL CONTROL AND THE SOCIAL ADJUSTMENTS OF CHILDREN[1]

RALPH M. STOGDILL
Wayne County Training School, Northville, Michigan

THE importance of the problem of parental attitudes was perhaps first recognized in the psychological clinic, where it was early noted that the behavior of the parent may be of the utmost significance in the causation of behavior problems in the child. Healy was one of the first to publish comprehensive case histories showing that both delinquency and neurotic behavior are more likely to result when home influences are unfavorable. These early clinical observations are now being verified by means of objective measurements in the field of parent-child relationships.

In the present experiment attitude questionnaires were constructed for the measurement of two types of parental attitudes which have been found to be significant in determining the behavior adjustments of children. Scale A was used to measure attitudes toward the parental control of children. Scale B was devised to reveal whether adults approve of an introvert or an extrovert social adjustment on the part of the child. Scale A consisted of 70 statements to be rated on a seven step scale from *extreme agreement* through *neutral,* to *extreme disagreement.* The subjects were instructed to indicate their responses by drawing a circle around one of the seven symbols after each statement. The significance of the symbols is as follows: large A indicates "very strongly agree"; middle A, "strongly agree"; little A, "agree"; the question mark, "neutral, undecided"; little D, "disagree"; middle D, "strongly disagree"; and large D, "very strongly disagree." Examples of the items are shown below:

[1] This investigation was conducted under the direction of Dr. Henry H. Goddard, Head of the Department of Abnormal and Clinical Psychology, The Ohio State University.

Parents should watch their children to see that no harm comes to them ..... A A A ? D D D

Parents should not try to control the behavior of their children ..... A A A ? D D D

A child should always believe what his parents tell him ..... A A A ? D D D

Children should be allowed to do as they please ..... A A A ? D D D

Children should have considerable supervision in all their activities ..... A A A ? D D D

Scale B consisted of 99 items to be rated in the same manner from *extreme agreement* to *extreme disagreement*. Examples of the items:

A child should be very polite ..... A A A ? D D D

A child should feel bashful about putting himself forward ..... A A A ? D D D

A child should be dissatisfied unless he is the leader in his group ..... A A A ? D D D

A child should worry over humiliating experiences ..... A A A ? D D D

A child should have complete confidence in his own ability ..... A A A ? D D D

The two scales were administered to psychologists, students and parents. The psychologists were members of the American Psychological Association who were specializing in child psychology or mental hygiene. The psychologists who rated each scale were divided into approximately equal groups for purposes of comparison. The students were obtained from classes in psychology. The parents were all mothers attending study groups or Parent-Teacher Association meetings. These mothers came from homes well above average in socio-economic status.

RESULTS OF SCALE A—ATTITUDES TOWARD THE PARENTAL CONTROL OF CHILDREN

Each item of Scale A was scored on an arbitrary (0 to 6) scale so that those individuals who endorse control over chil-

dren or disapprove of freedom receive low scores, while those who approve of freedom and disapprove of control receive high scores.[2] Any individual who "very strongly agrees" with an item endorsing control or "very strongly disagrees" with an item endorsing freedom receives a score of zero on that item; while a score of 6 is received if the subject "very strongly agrees" with an item endorsing freedom or "very strongly disagrees" with an item endorsing control. Intermediate values are assigned to positions between zero and six. A scale value of 3 is assigned to the neutral position on each item. Only 68 items of Scale A were used in computing total scores. The possible range of scores, thus is 0 to 408. The neutral score is 204. The use of the arbitrary scoring method appears justified since the correlation coefficients obtained between the arbitrary scores and a more elaborately weighted "sigma" scoring system range from 0.978 (P.E. = .005) to 0.995 (P.E. = .003).

Examination of Table I shows that the average scores for psychologists are considerably above neutral, and those for parents are below neutral. The average score for Psychologists II is 256.67; for Students II, 213.13; for Parents II, 182.61. Psychologists approve of freedom for children and disapprove of control. Students in Elementary Psychology (student groups II and III) rate only slightly above neutral; that is, they endorse very slight freedom for children. Student group I (a class in mental hygiene) approves of greater freedom for children than do the students in elementary psychology. Parents rate, in general, toward the control end of the scale. An interesting contradiction appears in the ratings of the parent groups. Whereas psychologists approve of freedom and disapprove of control, parents approve of both

[2] Twenty judges (seventeen faculty members and three graduate students of the Psychology Department of the Ohio State University) rated each item of Scale A as to whether the statement endorsed freedom for children or control over children. In addition, these judges rated each item of Scale B as to whether the statement represented an introvert or an extrovert form of behavior on the part of the child.

TABLE I

Means, Standard Deviations and Reliability Coefficients of Psychologists, Students and Parents on the "Attitudes Toward the Parental Control of Children" Questionnaire

	PSYCHOLOGISTS			STUDENTS			PARENTS	
	I	II	Total	I	II	III	I	II
N	36	39	75	56	92	66	33	46
MEAN	242.78	256.67	255.00	231.78	213.13	211.21	188.33	182.61
S.D.	43.24	31.08	37.13	36.81	29.27	38.13	35.81	30.04
r (split half)			.877	.834	.689	.833	.530	.494
P.E.			.018	.028	.037	.026	.086	.075
r (corrected)			.934	.909	.816	.909	.693	.661

control and freedom. Parents endorse such strong control, however, as to render the approved freedom only nominal.

There is considerable overlapping between all the groups. Statistically reliable differences are found between the means of the psychologists and those of all the student and parent groups. Critical ratios of the differences between the means of "Psychologists Total" and the different student groups range from 3.6 to 8.1, while those for "Psychologists Total" and the two parent groups are 8.9 and 11.7. Critical ratios of the differences between the means of the student and parent groups range from 2.9 to 7.4.

Reliability coefficients, computed by the split half method, and corrected by the Spearman-Brown formula, range from 0.66 to 0.93. These reliability coefficients, especially those for psychologists and students, are as large as are usually found in attitude measurements. Those obtained for parents are not as high, due probably to the unfamiliarity of these groups with such rating devices.

An average score of 236.13 was obtained when Scale A was rated at home by 40 advanced students in a class in abnormal psychology. When the scale was rated in class by this same group ten days later an average score of 236.63 was obtained. The correlation between the two sets of scores was found to be 0.947 (P.E. $\pm$ .011).

An average score of 222.00 was obtained when the scale was rated (by 25 advanced students in a class in personality) as it applied to six-year-old children, and an average score of 252.00 was obtained when the scale was rated again by the same group as it applied to fourteen-year-old children. This indicated that these students approved of a greater degree of freedom for the older children than they do for the younger ones.

That the scale does not measure intelligence is shown by a correlation coefficient of 0.005 (P.E. $\pm$ .091) obtained between the scores of 54 students from Group I and their percentile scores on the Ohio State Psychological Test.

Although no statistically reliable differences in attitude scores were found when the students (N = 158) in elementary

psychology were compared as to age, sex, class rank, college, major subject and church affiliation, several points are worth mentioning. Women students approve of greater freedom for children than the men do. Students under 17 years of age make more liberal scores than the average for the group, while students over 23 years of age make scores below the average for the group. Sophomores rate somewhat more conservatively and Seniors somewhat more liberally than average. Arts, Commerce, and Education College students make approximately the same scores. Several students majoring in psychology and home economics rate more liberally (*i.e.*, approve of greater freedom) than average, while those students majoring in engineering, medicine and physical sciences rate somewhat more conservatively than the average for the total group. Students affiliated with the Jewish and Protestant churches make average scores, while those affiliated with the Catholic Church endorse stricter control than average. Those who are affiliated with no church make considerably more liberal scores than the average. No difference appears between the average scores of those students who believe that they had "overstrict" home discipline and those who feel that they have had "too lax" home discipline. On the average, students who believe that they were "severely punished" as children endorse strict control, while both those students who "resent having been punished" and those who believe that their parents were "too moralistic" make more liberal scores than the average for the group.

No differences appear when psychologists are classified as to sex, marital status, and having or not having children of their own. When psychologists are classified as to geographical location, it is found that those living in the Middle-West and the East make more liberal scores than those living in the North-East, Far West, or South. Those psychologists who received their B.A. degree between 1890 and 1910 make more conservative scores than those who received their degrees before 1890 or later than 1910.

A study of a small group of parents (N = 15) whose children were held on probation by a Juvenile Court shows that these parents endorse much stricter control over children than any of the other groups studied. Although these findings are, in general, consistent with the results of previous investigations in parental attitudes and parent-child relationships, they should be interpreted with caution due to the small number of cases. The average score of this group is 136.6. These parents test about average on the Ohio Literacy Test and on the Clark-Thurstone Personality Schedule. The same tests were also given to the delinquent children (N = 15) of these parents. The children proved to be slightly more intelligent than their parents and more liberal in their attitudes as well as somewhat better adjusted emotionally. The average score on Scale A is 154.4 for these children. The Parental Control Scale is not correlated with intelligence, but there is a tendency in this group of parents and children for those who are better adjusted emotionally to make comparatively more liberal attitude scores and for those who are poorly adjusted emotionally to endorse extremely strict control over children. The attitude scores of *over-strict* and *over-lax* parents are very similar. The highly neurotic, but intelligent, *over-protective* parents and the *neglectful* parents of inferior intelligence both endorse more strict control over children than do any of the other parents in this group. The neglected children, although considerably less intelligent than the other children studied, are, nevertheless, better adjusted emotionally and are more liberal in their attitudes. The children with lax home discipline are more conservative in their attitudes than are those in the other three groups.

RESULTS OF SCALE B—ATTITUDES TOWARD CHILD BEHAVIOR

Each item of Scale B was scored on a 0 to 6 scale. The zero end of the scale was so placed that those individuals who approved of extrovert, and disapprove of introvert behavior

make low scores. The assigning of zero values to one end of the scale rather than another was purely an arbitrary procedure. It was intended that individuals with more liberal attitudes toward child behavior would make lower scores on this basis of scoring.

The results of Scale B are shown in Table II. The average scores for psychologists and students in the Mental Hygiene class are almost identical. The average score for psychologists is 221.84; for students 222.75, and for parents 260.52. Psychologists and students both approve of extrovert and disapprove of introvert behavior to a much greater extent than parents do. Statistically significant differences are found between the average score for parents and the average for Psychologists I, Pspchologists II, and students. The critical ratios between the mean for parents and the means of the other three groups are (in the order named) 7.7, 7.5, and 7.1. Corrected reliability coefficients for Scale B range from 0.64 to 0.90. This is somewhat lower than was found for the Parental Control Scale.

When the psychologists are classified as to sex, age, marital status, geographical location and decade of receiving B.A. or other equivalent degree the same trends are shown as were obtained with Scale A. Those psychologists receiving their B.A. degrees between 1890 and 1910 make more conservative scores; and those living in the Middle West and East make more liberal scores. In view of these results it would appear that the scale measures somewhat the same factor or factors measured by the Parental Control Scale. Further evidence that those individuals who express liberal (or conservative) attitudes toward children on the Parental Control Scale are likely to express the same kind of attitudes on the Attitudes Toward Child Behavior Scale is available in the form of correlation coefficients computed between the scores made on the two questionnaires by the groups that rated both. These correlations, when corrected for attenuation by Spearman's formula $\left(r = \frac{r_{12}}{\sqrt{r_{11}}\sqrt{r_{II2}}}\right)$, are $-.94$ for psychologists, $-.56$

for parents and −.49 for students. Those individuals who endorse freedom for children also approve of extrovert behavior, while those who endorse control over children tend to approve of introvert behavior. These results would be expected on *a priori* grounds. Parents who endorse strong control would be expected to approve of quiet, unaggressive children.

TABLE II

Means, Standard Deviations and Reliability Coefficients of Psychologists, Students and Parents on the "Attitudes Toward Child Behavior" Questionnaire

	PSYCHOLOGISTS			STUDENTS	PARENTS
	I	II	TOTAL		
N	46	41	87	58	38
MEAN	223.69	219.75	221.84	222.75	260.52
S.D.	26.15	29.26	27.73	33.82	18.05
r (split half	——	——	.730	.820	.473
P.E.	——	——	.033	.030	.090
r (corrected)	——	——	.843	.901	.642

CONCLUSIONS

Psychologists, on the average, approve of freedom for children and disapprove of strong external control. They approve of an extrovert social adjustment in a child, and disapprove of extremely introvert behavior. Students, in general, take a neutral stand on these questions. Parents, on the average, express the belief that children should have some freedom, but at the same time, they insist that children should be kept under control so strict as to render this freedom practically inoperative. As compared with psychologists, parents approve of introvert, and disapprove of extrovert social behavior.

Those individuals who endorse strong parental control over the child also approve of his making an introvert social adjustment while those who sanction freedom for children approve of extrovert social behavior.

SAMPLING THEORY AS USED IN THE DETERMINATION OF PSYCHOLOGICAL TRENDS IN VOLUME MAIL

MARY M. ROOS
Colorado College

SECTION I

IT is frequently of value to know the general trends of a large block of mail, or the significance of circular or questionnaire returns, or the size and type of sample necessary to be reasonably certain of rather definite reactions to some type of advertising. These reactions and others which have to do with the psychological responses of large groups of people to definite stimuli can be measured with a fair degree of accuracy provided the psychologist is careful in his analysis of what constitutes a trend and provided he uses the technique of sampling theory intelligently.

The practical experience basic to this paper was obtained while analyzing Federal Emergency Relief Administration mail, Federal Civil Works Administration mail, and in one instance, all White House mail in Washington, D. C., during the year 1933–34. However, the data on which the analyses were based are largely confidential in nature and hence in general the form and method only of their presentation can be given here. Only in cases where the data are obviously not confidential do the percentages given in the illustrative analyses represent actual situations. In cases where data do not represent actual situations they have been obtained from factual material by a formula which the author will publish whenever its revelation is proper.

When one is confronted with a vast volume of mail, before any analysis can be attempted it is necessary to determine what types of psychological trends are being sought. Generally one has in mind a broad outline of the types of trends

for which he is looking. For instance, those sought in the recent Federal mail analyses measured reactions toward various phases of the New Deal. These nation-wide reactions were pronounced when some new policy was announced such as the "Federal Aid to Students" policy, or some new alphabetical agency set up such as the Federal Civil Works Administration. Then again there may be some specific trends which challenge observation as there were in a special political and legislative policy mail analysis. In this last instance the strength of the psychological trends for and against such Congressional bills as the Soldiers' Bonus bill, the Fletcher-Rayburn bill, and also the sentiment regarding Old Age Pension legislation, was measured and analyzed. In this analysis and also in all others made, the occupational groups of the people who reacted and the sections of the country from which they came were noted.

The sources of stimuli for the psychological trends found in the Federal mail are many, *i.e.,* radio speeches, newspaper dispatches, editorials, new governmental policies, and so forth. The public response to President Roosevelt's radio speeches was almost immediate. Indeed, usually within three days the psychological reactions throughout the country to any new government policy whether announced over the radio or through the newspapers began to be evident in the Federal mail. The quick response to the following administrative policy may be cited as an instance of the sensitivity of the public mind. On January 18, 1934, Harry L. Hopkins stated publicly that, effective February 15, 1934, the Federal Civil Works Administration activities would be materially curtailed unless adequate appropriations were made by the legislative branch of the government. In the mail analysis for the week ending January 27, 1934, 21.4 per cent of all communications requested continuation of the Federal Civil Works activities, and 12.6 per cent warned of dire social crises if the program were discontinued.

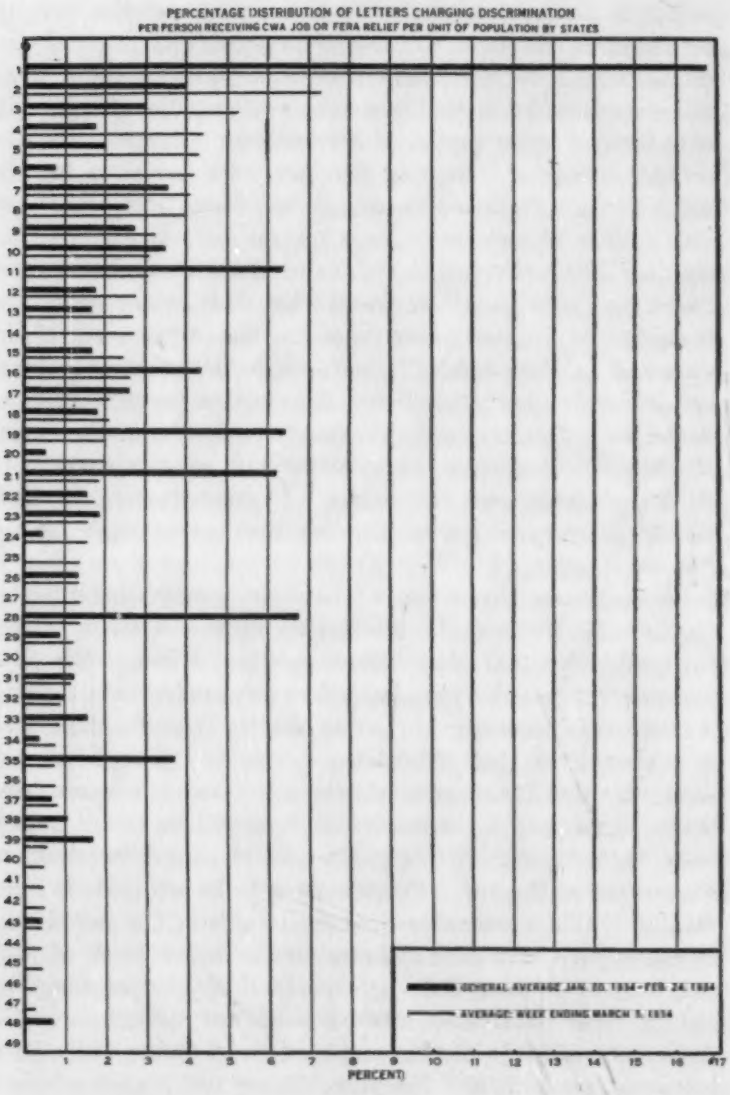
Following this announcement by Mr. Hopkins the newspapers stressed two items, (1) charges of graft regarding the misuse of Federal Civil Works Administration Funds, and (2) an announcement that an appropriation of $950,000,000 was being requested to carry on the Relief and Civil Works programs. Both of these items were forcefully reflected in the mail for the week ending February 2, 1934, in which only 4.2 per cent requested continuation of the Civil Works activities and only 1.6 per cent warned of dire social crises if such activities were not continued beyond February 15, 1934 (the curtailment date mentioned by Hopkins). The two factors had operated in such a way as in the first instance (newspaper publicity to charges of graft regarding the misuse of public funds) to shake the confidence of one group of people in the manner in which public funds were being used, and in the second instance (the public announcement that $950,000,000 was being requested from Congress for Relief and Civil Works programs) to quiet the fears of those who were doing or needing the work provided by the Civil Works program.

As indicated above the newspapers were also a constant and important source of the psychological trends found in the mail. When the newspapers began publicizing the graft allegations given above, the number of letters charging some form of graft greatly increased. These trends seem rather akin to mass psychology. If the newspapers publicized graft charges the consensus of opinion was quickly atune to look for graft and this atuning with what was considered the popular mind evidenced itself with great rapidity in charges of graft which in a few instances upon investigation proved to be real, but which generally were found to be imaginary.

The Federal mail analyses were used for many practical purposes. A typical example of such a use follows. However, because of the confidential nature of the content, here only the form of the data can be given. Also, because of the obvious derogatory character of the data given on the chart, *numbers* are substituted for the names of the states.

In this instance, it was desired to know at a particular period in the spring of 1934 what sections of the country were making the most charges of personal and political discriminations. It was further desired to know how these charges compared with the averages of similar charges per state for the entire period of the analyses. Accordingly, the average charge of discrimination per week per state for the entire period of the mail analyses was found and contrasted with similar immediate findings for the periods under investigation. The chart which follows illustrates clearly just where the "sore spots" were and what their extent was so far as charges of discrimination regarding the carrying out of the state and local Federal Civil Works Administration policies and Federal Emergency Relief Administration policies were concerned. For example, it is simple to determine the extent of such dissatisfaction in state no. 10 as compared with all other states, and the extent of dissatisfaction in state no. 10 as compared with its own average of such former charges.

Regardless of the extent of familiarity with the types of trends sought, the challenge of any block of mail is to find the type and extent of all trends existing in it. Hence, it is perhaps well to point out that before any analysis of any kind is made, it is necessary to read at random from the mail until it is possible to form a trial key. This key must have such amplitude that for an adequate sample chosen at random from the block of mail to be analyzed there will be found practically no motive for writing a letter which cannot be tabulated somewhere in the key. For a sample to be adequate it must exhibit within reasonable preassigned limits the percentage of the motive in question shown in the entire block of mail which is to be analyzed. Obviously this involves sampling theory. The statistical method used for determining the degree of adequacy of the sample is given under "2" of the memorandum entitled "Sampling Theory and Mail Analysis," reproduced in Section II.


PERCENTAGE DISTRIBUTION OF LETTERS CHARGING DISCRIMINATION
PER PERSON RECEIVING CWA JOB OR FERA RELIEF PER UNIT OF POPULATION BY STATES
GENERAL AVERAGE JAN. 20, 1934 - FEB. 24, 1934
AVERAGE: WEEK ENDING MARCH 3, 1934
PERCENT)

After the key has been found adequate, each reader is given a copy of it, and every possible bit of useful information relative to the correspondent or his message is tabulated on the reader's chart key. In the instances of the Federal mail analyses the reader's chart keys were later compiled into the same general types of analyses as are illustrated by the hypothetical ones given at the end of this paper.

Since nothing is left out of these analyses and because an individual may have several different expressed motives for writing a letter the motivation columns in analysis (II) and the subject columns in analysis (I) must be read thus: "A given per cent of all letters had to do with a given motive or subject." Since most letters are written because of more than one motive or purpose the motivation and subject columns cannot total 100 per cent. The figure basic to the per cent is the total number of letters analyzed and not the total number of motives expressed. All other columns do, however, total 100 per cent.

The actual statistical technique involved in volume mail analysis is of a rather elementary nature. Inasmuch as the following memorandum contains in detail the methods used in applying sampling theory to volume mail, it is reproduced here essentially as it was presented to President Roosevelt anticipatory to analyzing all White House mail.[1]

Section II

SAMPLING THEORY AND MAIL ANALYSIS

Sampling theory, if it is thought of at all, by most people is thought of as a highly theoretical branch of statistics. The recent Mail Analyses of the Correspondence Division of the Federal Emergency Relief Administration and the Federal Civil Works Administration are examples of the simplicity and utility of sampling theory in a practical situation.

It is one thing to have a vast volume of mail continuously being received and answered by our President and those who

[1] It was on the basis of this memorandum that the President requested an analysis of all White House mail.

work with him, and it is quite another thing to know in one picture who wrote the letters and why they were written. In the first instance the letters are answered letter by letter. Only a detailed fragmentary impression of their general content remains in the minds of each of several dictators and readers. The people of the United States are speaking individually to their President and to his coworkers but their voices cannot be heard. As far as is practicable their individual needs are met, but at the end of the week there is no picture of what these thousands of people have been trying to tell their President and those whom he has chosen to help him. It is the realization of this need for a complete picture of immediately expressed desires of American citizens which first inspired the Mail Analyses.

Because of the great cost, both in money and time, for reading and analyzing each letter separately, a very practical branch of pure statistical theory, known as sampling theory, was used to determine the three following entities:

1. The validity of the measurements of the analysis.
2. The extent of an adequate sample.
3. The degree of temporariness or permanency of the trends.

No attempt will be made in this survey to give a complete mathematical presentation of all the applications of the sampling theory which make these analyses possible. However, for the purpose of clarity a simple illustration of the above three methods of procedure follows:

1. *A Method of Determining the Validity of the Measurements of the Analysis*

Suppose first that it is desired to know the validity of the measurements of some one percentage in a sample analysis as related to the total sample of letters being analyzed. To solve this problem let,

N = the number in the sample,

$X = \frac{\%}{100}$ be the true proportion of all letters containing a particular attribute y, for instance, the per cent of mail requesting direct relief,

σ = standard deviation of X in sets of N cases.

Then, as is well-known,[2] $\sigma/N = \sqrt{X(1-X)/N}$. An examination of a table of the normal probability function shows that only 37 out of every 100 samples of N letters will not yield percentages of y within the range $X = 2\ \sigma/N$. By increasing the multiplier of σ/N the number of cases outside the range can be reduced to as small a value as desired; but for the purpose of adequately checked mail analyses 2 has been found satisfactory.

2. *A Method for Determining the Extent of an Adequate Sample*

One of the methods for deciding the extent of an adequate sample is based on the formula just outlined for determining the validity of the measurements of the items within the analyses. In fact, one merely needs to choose σ/N arbitrarily so that the range $X \pm 2\ \sigma/N$ is satisfactorily small, and from the formula

$$\sigma = \sqrt{NX(1-X)} \text{ calculate } N.$$

3. *A Method of Determining the Degree of Temporariness or Permanency of the Trends*

Suppose that it is desired to determine whether or not the per cent of individuals who complained of inability to obtain jobs during the period of five mail analyses come from a homogeneous group. In other words, suppose, for instance, that it is desired to determine whether or not the variation of complaints regarding inability to obtain jobs (1) represents a definite trend within a certain group of people or (2) is a temporary shift which is continuously influenced by ex-

[2] See, for example, H. L. Rietz, "Handbook of Statistics," 1934, p. 72.

traneous factors. The second instance would obviously be due to an error in sampling the population, that is, failure of the mail to adequately represent the true state of public reaction.

The best estimate X of the percentage complaining in the total population is given by $X = 2239/4975 = .450$. On this basis the expected numbers complaining at each of the dates are 338, 440, 543, 457, and 461, respectively, obtained by multiplying the number in each sample by .450.

Pearson's X^2 (chi square) function may be used, to test whether the variations between actual and theoretical values can be attributed to chance. Chi square is readily computed since the contribution of each observed value to it is equal to the ratio of the square of the difference between the theoretical and actual values to the theoretical value. Thus in this example, the contribution of 497 to chi square is $(497–338)^2/338 = 159 \times 159/338 = 74.8$. The remaining values for those complaining contribute 17.6, 23.1, 0.2, and 5.6, respectively to chi square. The number not complaining contribute 61.2, 14.4, 18.9, 0.2, and 4.6, respectively. The total contribution to chi square is then the sum of 17.6, 74.8, 23.1, 0.2, 5.6, 61.2, 14.4, 18.9, 0.2, 4.6, = 220.6. For a two column five row contingency table, such as obtained here, the tables for chi square must be entered for $N = (S-1)\ (2-1) = 4$.[3]

A reference to Table III of R. A. Fisher, *Statistical Methods for Research Workers*, 5th edition, shows that the probability, P of getting such a high value of chi square from a homogeneous population is very much less than .01.[4] Consequently, the data were definitely non-homogeneous, that is, trends were in evidence.

[3] Considering the tabulation of actual and theoretical numbers as a two-fold rectangular array or contingency table, the N, the number of degrees of freedom, is equal to the number of rows less 1, times the number of columns less 1.

[4] For P less than .01, we can be quite certain that there is no homogeneity in the set of numbers studied, and for P between .01 and .05 there is a fair probability that this is true.

Complaints of Inability to Obtain Jobs

TIME	ACTUAL				THEORETICAL		ACTUAL MINUS THEORETICAL	
	Total Letters Received	Number in Sample	Number Complaining	Number not Complaining	Number Complaining	Number not Complaining	Complaining	Not Complaining
Jan. 20	21,436	751	497	254	338	413	159	−159
Jan. 27	28,000	978	352	626	440	538	−98	98
Feb. 3	16,000	1,206	431	775	543	663	−112	112
Feb. 10	14,500	1,016	447	569	457	559	−10	10
Feb. 24	9,648	1,024	512	512	461	563	51	−51
Totals		4,975	2,239	2,736	2,239	2,736		

Clearly the test given above can be used also to compare percentages for successive weeks. Suppose, for instance, that it is desired to test the two weeks February 10 and February 24. The best estimate of X here is $(447+512)/(1016+1024) = 959/2030 = .4724$. The theoretical values of the numbers complaining are $.4724\ (1016) = 480$, and $.4724\ (1024) = 484$, and for those not complaining 536 and 540 respectively. The contributions to chi square are 2.3, 1.5, 2.0 and 1.5, respectively, and $X^2 = 2.3 + 1.5 + 2.0 + 1.5 = 7.3$. In this example the table has dimension 2×2 so that $N = (2-1)(2-1) = 1$. Again P is less than 0.1 so that a definite trend is in evidence. On the other hand the data for the weeks January 27 and February 3 may be regarded as homogeneous. Here the theoretical numbers complaining are 351 and 432, and the theoretical numbers not complaining 627 and 774. The quantity $X^2 = .003$, and since $N = (2-1)\ (2-1) = 1$, $P = .92$. Thus, there are only 8 chances in 100 that the material is not homogeneous.

These results show clearly that the percentages complaining of inability to obtain jobs vary so widely that the total number of letters for all weeks cannot be regarded as having been drawn from the same population. Indeed, with the exception of the period from January 27 to February 3, in no two consecutive weeks were the percentages sufficiently close together to be consistent with an assumption of a homogeneous population. Hence, the conclusion is that the percentages studied follow a very definite trend during the period, reaching the bottom of a sharply declining movement between January 27 and February 23, and then increasing steadily at a significant rate during the rest of February.

The foregoing figures are largely self-explanatory. The sets of figures, "contributions to chi squared" are obtained for each space by dividing the square of the difference between the actual and theoretical numbers for that space by the theoretical number. Of course, the actual numbers were obtained by taking the percentage indicated in the number in the sample.

A further statistical analysis of the geographical distributions of letters in the mail analyses proved interesting. First, the percentage of letters written per state was determined, and then the percentage of total population per state was found. On the basis of these two percentages the geographical distribution of letters received per unit of population was calculated. Using the percentage of persons receiving either CWA jobs or FERA relief per state as a base the percentage of letters received per person accepting CWA jobs or FERA relief per unit of population per state was next determined. To put the matter simply, for the week ending March 3, 1934, 5.5 per cent of all letters received and handled by the Correspondence Division of the Federal Emergency Relief Administration and the Federal Civil Works Administration were from the state of Illinois. On the basis of a common unit of population for the United States, 1.5 per cent were from the State of Illinois. And on the basis of a common unit of population receiving CWA jobs or FERA relief within the United States, 8.4 per cent were from Illinois. These last two percentages are highly significant because they are ratios which adequately measure the extent of people writing.[5]

[5] The mathematics underlying the first determination is as follows:

Let X_i = Number of letters received in ith state.
Let P_i = Number of people in ith state.
Let Z_i = % of letters received in ith state.
Let Y_i = % of population in ith state.

then:

$$\frac{Z_i/Y_i}{\Sigma Z_i/Y_i} = \frac{\frac{X_i}{\Sigma X_j}\Big/\frac{P_i}{\Sigma P_j}}{\Sigma\frac{X_i}{\Sigma X_j}\Big/\frac{P_i}{\Sigma P_j}} = \frac{X_i/P_i}{\Sigma X_i/P_i}$$

which last is the per cent of letters received per person in the ith state. The ratio remains the same whenever P_i is multiplied by any number, so that in reality it is merely necessary to refer to this quantity as the per cent of letters received per unit of population, without specifying the size of the unit.

By similar mathematical analysis the percentage of letters received per person receiving CWA jobs or FERA relief per unit by state can be obtained.

Section III

The two following analyses are given as demonstrations of a possible form and method of presentation of a mail analysis. In no instance do the percentages given represent actual situations.

ANALYSIS I

Analysis of all President X's Mail Save that Which was Sorted for Special Attention for the Period March 23 Through March 27, 19—.

The following analysis of 2,544 letters represents a random sample of approximately 18,500 letters received by President X during the period March 23 through March 27, 193–. This sample represents all Presidential mail for this period, save that which was sorted for special attention.

I. *Analysis of Geographical Distribution*

GEOGRAPHICAL DISTRIBUTION OF LETTERS RECEIVED		GEOGRAPHICAL DISTRIBUTION OF LETTERS RECEIVED PER UNIT OF POPULATION	
Illinois	16.8	Illinois	6.0
New York	9.9	Delaware	5.4
Pennsylvania	6.5	Colorado	5.3
Missouri	5.2	Wyoming	4.5
California	4.2	Missouri	3.8
New Jersey	3.8	Minnesota	3.4
Minnesota	3.3	Mississippi	3.4
Ohio	3.1	Montana	3.2
Michigan	3.1	South Carolina	3.2
Indiana	2.8	Iowa	2.8
Wisconsin	2.7	Wisconsin	2.7
Iowa	2.7	New Jersey	2.4
Texas	2.7	Indiana	2.5
Mississippi	2.3	Kansas	2.3
North Carolina	2.1	New Mexico	2.2
South Carolina	2.1	New York	2.1
Colorado	1.8	California	2.1
Oklahoma	1.7	Oklahoma	2.1
Tennessee	1.9	Washington	2.1
Alabama	1.6	West Virginia	2.1

Kentucky	1.7	Tennessee	1.8
Kansas	1.4	Florida	1.7
Massachusetts	1.6	Nebraska	1.7
Georgia	1.2	Pennsylvania	1.7
Virginia	1.4	South Dakota	1.7
West Virginia	1.2	Alabama	1.6
Washington	1.1	Kentucky	1.6
Florida	1.1	Michigan	1.6
Connecticut	.8	North Carolina	1.6
Louisiana	.8	Oregon	1.6
Nebraska	.8	Utah	1.5
Maryland	.7	Connecticut	1.4
Arkansas	.7	North Dakota	1.4
Montana	.7	Rhode Island	1.4
Oregon	.7	Virginia	1.3
Delaware	.6	Maryland	1.3
South Dakota	.6	Ohio	1.3
North Dakota	.6	Georgia	1.2
Rhode Island	.5	Louisiana	1.2
Wyoming	.4	Texas	1.1
New Mexico	.4	Idaho	1.1
Utah	.3	Dist. of Columbia	1.1
Dist. of Columbia	.3	Arkansas	1.0
Idaho	.3	Massachusetts	1.0
Maine	.2	Maine	.9
Arizona	.2	Arizona	.8
New Hampshire	.2	New Hampshire	.8
Nevada	.1	Nevada	.4
Vermont	.1	Vermont	.3

II. *Occupational Analysis*

46.2% of the letters were from laborers.
17.1% were from business men.
15.2% were from farmers.
13.8% were from clerical workers.
3.2% were from professional people.
1.8% were from groups.
1.3% were from children.
.8% were from students.
.6% were from Government officials.

III. *Sex of Writer:* 73.0% were written by men; 27.0% were written by women.

IV. *Racial Analysis:* 97.7% were written by members of the white race; 2.3% were written by members of the colored race.

V. *Types of Communication Analysis*: 39.4% fair; 19.2% poor; 19.3% excellent; 22.1% propaganda form letters.

VI. *Addressee*

A. Primary Addressee

100% of the letters were addressed to President X.

B. Secondary Addressee

Per cent	
12.6	Department of Labor
12.4	Post Office Department
11.6	Home Loan Bureau
11.1	Farm Credit Administration
9.7	Federal Emergency Relief Administration
8.8	Federal Trade Commission
6.2	Treasury Department
5.5	Veterans' Bureau
4.4	National Recovery Administration
3.7	Civil Works Administration
2.6	Department of Justice
2.1	Public Works Administration
1.6	Reconstruction Finance Corporation
1.3	Department of Agriculture
1.3	Civilian Conservation Camps
.6	Department of Commerce
.5	Interstate Commerce Commission
.4	Department of Navy
.3	Department of Interior
.2	Department of War
.1	Tennessee Valley Administration
.1	Alcohol Control Board
.1	Agricultural Adjustment Administration

VII. *Subject:* 61.3% Personal; 34.7% Federal; 9.0% Local; 1.7% International; 1.1% State.

VIII. *Motivation*

A. Satisfaction

18.7% contained personal praise for President X.

11.5% were pledges of faith in and loyalty to the New Deal.
10.8% were praise for the personnel of the X Administration.
1.1% were expressions of gratitude for aid received.

B. Suggestion

9.1% contained suggestions that the President endorse some pending legislative bill.
5.5% contained suggestions for new administrative policies.
4.6% contained suggestions for improvement of present administrative poliices.
1.8% suggested new Federal projects or plans for ending the depression.
1.6% made suggestions regarding the improvements of present administrative machinery.

C. Request

Per cent

18.4 requested immediate relief.
15.2 employment.
11.4 requested loans for homes.
7.3 requested loans to save their farms.
7.2 requested personal endorsement from the President.
7.1 requested Federal investigation.
6.4 requested veterans' aid
4.1 requested the President's advice.
2.8 requested pensions.
2.6 requested general information regarding the Federal Government and its policies.
2.4 requested money to buy a farm.
2.4 requested higher wages and shorter hours.
2.1 requested information regarding departmental policy.
2.1 requested a personal loan.
1.7 requested a loan to provision a farm.
1.1 requested a loan in order that they might save their business.
0.6 requested autographs, photographs and gifts from the President.

0.5 requested information regarding the President's policy.
0.4 requested aid at Warm Springs, Georgia (from cripples).
0.4 requested that farm produce prices be raised.
0.3 made general requests.
0.3 were general requests for education aid.
0.3 requested aid in securing citizenship papers.
0.2 requested lower commodity prices.
0.1 requested payment of bonds or currency in gold when that statement was made on the bonds of currency.

D. Protest

Per cent

10.2 protested the enactment of a Federal Legislative bill.
8.8 protested personal treatment received from Government authorities.
4.6 charged discrimination.
3.8 made various charges regarding the inefficiencies and unfairness of the New Deal.
3.1 protested the activities and policies of the FERA.
2.5 protested the activities and policies of the NRA.
2.3 charged graft or the misuse of Government funds.
1.8 protested our present banking system and urged Federal control of same.
0.5 protested the practice of selling below the cost price.
0.4 protested the inconsistencies of the policy of the AAA and the needs of the present time.
0.3 protested the monetary policy of the New Deal.
0.2 protested the inadequacy of our present methods of crime control and urged strong Federal supervision of same.

E. Response

1.3% of all letters stated that they had written before.
0.5% of these had received a satisfactory response.
0.9% had either received no response at all or an unsatisfactory one.

Explanatory Notes:

Note 1: Percentages in Column "A" mean that:

16.8% of all letters analyzed came from the state of Illinois
9.8% of all letters analyzed came from the state of New York, etc.

These percentages are significant in that they point out in each state the degree of written expressed reaction to the President and to Federal Policies. Percentages in Column "B" measure the relative intensity of written expressed reaction per unit of population by state; we can know per unit of population that there are six times the expressed reaction to the President and to Federal policies in the state of Illinois as are in the state of Arkansas.

The unit of population may be any unit which we choose. In mathematical computations the actual unit of population vanishes and only percentages remain. *It is the difference in the percentages which is significant.* In the case of Illinois, which is responsible per unit of population for 6.0% of all letters analyzed, and Arkansas which is responsible for 1.0% of all letters analyzed, the significant feature is that the people of Illinois have expressed themselves six times as frequently as the people of Arkansas.

Note 2:

Percentages under the following headings total 100 per cent. Under these headings each letter can be analyzed only once.

I. Secondary Addressee
II. Occupational Analysis
III. Sex of Writer

IV. Racial Analysis
V. Type of Communication
VIII. Geographical Distribution

The subject and motivation of each letter may be analyzed objectively under several subheadings, *i.e.*, it may be both personal and international; satisfied with AAA activities but dissatisfied with NRA.

Percentages under heading (VI)—*Subject*—indicate that

61.3% of all letters analyzed had to do with the individual writer.
34.7% with the Federal Government.
9.1% with local conditions.
1.7% with international relations.
1.1% with state conditions.

These same letters had distinguishing points noted under Motivation.

Percentages under heading—*Motivation*—are based on the various motives in each letter analyzed. A man may be motivated by his admiration for the President; by his gratitude for a job received through government aid; he may want to suggest a new means of ending the depression, and desire a loan to save his farm.

ANALYSIS II

Federal Civil Works Administration

Federal Emergency Relief Administration

Summary of Mail Analysis for Week Ending March 10, 19—.

The following mail analysis of 1012 letters and telegrams represents a random sample of 9,208 Federal Civil Works Administration and Federal Emergency Relief Administration communications, including those referred from the White House, which were handled by the Correspondence Division during the period March 3, through March 10, 193–.

I. *Analysis of Geographical Distribution*

	GEOGRAPHICAL DISTRIBUTION OF LETTERS RECEIVED	GEOGRAPHICAL DISTRIBUTION OF LETTERS RECEIVED PER UNIT OF POPULATION	PERCENTAGE LETTERS RECEIVED PER PERSON RECEIVING CWA JOBS OR FERA RELIEF PER UNIT OF POPULATION BY STATE
	Per cent	*Per cent*	*Per cent*
Illinois	10.1	2.7	15.0
Georgia	5.5	8.6	7.0
New York	6,6	1.0	6.5
Michigan	4.6	2.0	5.5
North Carolina	4.0	2.6	4.2
Kansas	3.8	4.4	4.0
Virginia	2.5	2.2	3.4
Ohio	4.2	1.3	3.2
Oregon	3.0	6.6	3.1
Missouri	2.2	1.2	3.0
Tennessee	2.1	1.7	3.0
Texas	2.7	1.0	2.8
New Jersey	2.8	1.4	2.5
Pennsylvania	4.0	0.8	2.4
Indiana	2.7	1.8	2.4
California	2.1	0.7	2.3
New Mexico	1.6	9.7	2.2
Louisiana	2.5	2.5	2.0
Oklahoma	2.0	1.7	1.8
Mississippi	2.4	2.6	1.7
Iowa	1.4	1.2	1.6
Alabama	3.7	3.0	2.3
Massachusetts	1.8	.8	1.8
Wisconsin	1.7	1.3	1.6
Kentucky	2.4	1.9	1.6
Washington	1.9	2.6	1.5
Arkansas	1.8	2.1	1.3
West Virginia	2.4	2.9	1.1
South Carolina	1.5	1.8	.9
Florida	1.9	2.7	.9
Maine	.7	1.5	.8
Maryland	.9	1.2	.8
Arizona	.9	3.6	.7
Connecticut	.7	.9	.7
Dist. of Columbia	.6	2.3	.7

	GEOGRAPHICAL DISTRIBUTION OF LETTERS RECEIVED	GEOGRAPHICAL DISTRIBUTION OF LETTERS RECEIVED PER UNIT OF POPULATION	PERCENTAGE LETTERS RECEIVED PER PERSON RECEIVING CWA JOBS OR FERA RELIEF PER UNIT OF POPULATION BY STATE
	Per cent	*Per cent*	*Per cent*
Minnesota	.8	.7	.7
Montana	.7	2.7	.6
North Dakota	.8	2.1	.5
New Hampshire	.4	1.4	.4
Vermont	.3	1.3	.4
Colorado	.3	.5	.3
Nebraska	.3	3.2	.3
Rhode Island	.2	.5	.2
Utah	.2	.4	.2
Delaware	.2	.9	.0
Idaho	.0	.0	.0
Nevada	.0	.0	.0
South Dakota	.0	.0	.0
Wyoming	.0	.0	.0

II. *Occupational Analysis*

61.5% of the letters were from laborers.
9.7% of the letters were from clerical workers.
7.6% of the letters were from professional people.
7.3% of the letters were from children.
6.2% of the letters were from civic groups.
5.6% of the letters were from foreigners.
2.1% of the letters were from students.

III. *Racial Analysis:* 97.4% were from the white race; 2.6% were from the colored race.

IV. *Type of Communication Analysis*

A. Letters

37.6% fair; 31.7% poor; 27.1% excellent.

B. Telegrams

3.6% were telegrams.

V. *Motivation Analysis*

A. Praise

3.4% expressed praise for President Roosevelt and his Administration.

3.6% expressed gratitude for aid received.

B. Suggestions

3.1% gave general suggestions for bettering the Federal Government.

2.0% suggests new plans for ending the depression.

1.9% suggested new Civil Works projects.

C. Information

3.7% asked information regarding wages on Civil Works projects.

3.7% asked for general information.

D. Requests

28.5% requested direct relief.

23.2% requested help in obtaining a job.

13.7% requested general information.

5.6% requested educational assistance.

3.3% requested loans.

2.2% requested the President's advice on personal matters.

2.0% requested information on marginal land project.

1.3% requested aid in obtaining farm implements, seed, etc.

.3% requested increased CWA quotas.

E. Complaints

41.4% complained of inability to obtain, and need for, jobs.

27.8% protested the personal treatment received from the local Civil Works Administration.

15.0% protested the demobilization of CWA because of the great need of the workers.

13.5% protested the general inadequacy of the CWA and FERA for local needs.

13.1% protested the personal treatment rereceived from the local FERA.

12.6% charged local CWA or FERA discrimination of the following types:
1. 60.7% personal
2. 35.3% political
3. 8.1% negroes
4. 8.0% non-union
5. 2.3% union

10.4% protested that jobs or relief were given to those who did not need them.

7.5% complained of graft and the misuse of Government money.

6.7% were from veterans protesting the cutting of pensions, compensations, or lack of veterans' preference for CWA job selection.

6.3% protested the demobilization of the CWA because of the great economic waste which would be increased as a result of unfinished CWA projects.

5.5% protested the CWA and CCC salaries and FERA relief were inadequate to family needs.

1.1% were from farmers and legitimate business protesting competition of CWA.

.2% protested lack of occupational selection of personnel with regard to abilities.

.2% protested the practical exclusion of single men from work and relief.

Geographical Distribution of Letters Charging Discrimination

STATE NUMBER	GEOGRAPHICAL DISTRIBUTION OF LETTERS CHARGING DISCRIMINATION	GEOGRAPHICAL DISTRIBUTION OF LETTERS CHARGING DISCRIMINATION PER UNIT OF POPULATION BY STATE	PERCENTAGE LETTERS CHARGING DISCRIMINATION PER PERSON RECEIVING CWA JOBS OR FERA RELIEF PER UNIT OF POPULATION BY STATE
One	3.9	4.2	7.4
Two	3.9	3.7	6.8
Three	4.7	1.0	6.4
Four	28.5	10.4	5.9
Five	3.1	2.3	5.9
Six	4.7	4.1	5.8
Seven	3.1	1.5	4.8
Eight	3.9	2.7	4.8

STATE NUMBER	GEOGRAPHICAL DISTRIBUTION OF LETTERS CHARGING DISCRIMINATION	GEOGRAPHICAL DISTRIBUTION OF LETTERS CHARGING DISCRIMINATION PER UNIT OF POPULATION BY STATE	PERCENTAGE LETTERS CHARGING DISCRIMINATION PER PERSON RECEIVING CWA JOBS OR FERA RELIEF PER UNIT OF POPULATION BY STATE
Nine	4.7	7.1	4.6
Ten	3.1	18.1	4.5
Eleven	3.1	1.5	4.5
Twelve	3.9	6.8	4.2
Thirteen	3.1	1.3	3.3
Fourteen	2.4	1.6	3.0
Fifteen	1.6	1.8	3.0
Sixteen	1.6	.9	2.6
Seventeen	3.9	7.4	2.5
Eighteen	1.6	1.8	2.5
Nineteen	1.6	2.8	2.3
Twenty	1.6	1.4	2.3
Twenty-One	2.4	.7	2.1
Twenty-Two	1.6	1.8	2.0
Twenty-Three	1.6	1.5	2.0
Twenty-Four	1.6	2.3	1.6
Twenty-Five	.8	1.2	1.2
Twenty-Six	.8	4.5	1.0
Twenty-Seven	.8	.9	.9
Twenty-Eight	.8	.8	.7
Twenty-Nine	.8	.9	.7
Thirty	8.	3.0	.7
Thirty-One	.0	.0	.0
Thirty-Two	.0	.0	.0
Thirty-Three	.0	.0	.0
Thirty-Four	.0	.0	.0
Thirty-Five	.0	.0	.0
Thirty-Six	.0	.0	.0
Thirty-Seven	.0	.0	.0
Thirty-Eight	.0	.0	.0
Thirty-Nine	.0	.0	.0
Forty	.0	.0	.0
Forty-One	.0	.0	.0
Forty-Two	.0	.0	.0
Forty-Three	.0	.0	.0
Forty-Four	.0	.0	.0
Forty-Five	.0	.0	.0
Forty-Six	.0	.0	.0
Forty-Seven	.0	.0	.0
Forty-Eight	.0	.0	.0

A TABLE OF PERCENTILE EQUIVALENTS FOR EIGHT INTELLIGENCE TESTS FREQUENTLY USED WITH ADULTS[1, 2]

IRVING LORGE
Teachers College, Columbia University

THERE are many tests of intelligence available to the clinician, to the psychologist, to the educator, to the counselor, and to those others interested in, or using psychotechnics. The advantages of many intelligence tests, however, are obscured partly by a Babel of scores. There is a genuine need for tables giving the percentile values of many intelligence tests for transmutation. By means of adequate conversion tables some of the confusion attendant on the interpretation of intelligence scores, or on the assignment of levels, might be resolved. In the course of a long-time experiment in adult learning, a larger group of subjects was given many tests. From this group were selected eighty adults, ranging in age from twenty to seventy years, each of whom had taken *at least one form*[3] *of each of*

CAVD: The I. E. R. Intelligence Scale CAVD Levels M–Q, Forms 1, 2, 3, 4, and 5.

Thorndike: Thorndike Intelligence Examination for High School Graduates, Forms AA, AA, R; BB, BB, W; CC, CC, Y; EE, EE, DD; and FF, FF, V.

[1] Acknowledgment is hereby made of the services rendered by the personnel furnished by the Works Division, Emergency Relief Bureau of New York City on Project 89FB–125X.

[2] This study is part of a larger study in Interests, Attitudes, and Motives supported in part by a grant from the Columbia University Council for Research in the Social Sciences.

[3] If more than one form was taken, the average of the scores was used. Usually each subject took all five forms of each the CAVD and the Thorndike unless he or she was separated from the project.

Army Alpha: The United States Army Group Examination Alpha, Forms 5 and 7.

and each of

Wells: Revised Alpha Examinations (form 5).

Bregman: Revision of Army Alpha Examination.

Otis: S–A Tests of Mental Alertness. Higher Forms A and B (20 minutes).

Pressey:[4] Senior Verifying Test and Senior Classification Test.

Bureau VI:[4] Bureau VI Mental Alertness. Forms A and B.

The order in which the tests were usually administered was CAVD, Army Alpha, Bregman, Wells, Thorndike, Otis B, Pressey Classification, Bureau B, Pressey Verification, Otis A, and Bureau A. The time over which these tests were administered was seventeen months at the maximum and two months at the minimum.

From the scores of this sample, the first and ninety-ninth percentile and each semidecile were computed. These percentiles are reported in Table I. It is apparent that the eighty adults represent a superior sampling of the general adult population. For instance, the medians of this group are equivalent to the

thirtieth percentile of college graduates on the CAVD
ninety-sixth percentile of general adult population in the Army Alpha
fiftieth percentile of college students on the Otis B.

The tables, nevertheless, are of value in providing an approach to the solution of the problem of the use of many different tests in the development of educational and occupational standards. The table can be of value in the conversion of some of the tests more frequently used with adults.

It is essential, however, to note that the assumption of the equivalence of percentile values for the various tests is tantamount to regarding the intercorrelations among these tests as being uniformly unity.

[4] The two Pressey tests and the two forms of the Bureau VI were administered at the request of Dr. Walter V. Bingham who supplied the tests.

TABLE I

The score by specified percentiles, derived from the results of an identical adult population of eighty persons, each of whom took each of the following tests:

CAVD: I.E.R. Intelligence Scale CAVD, Levels M–Q
Army Alpha: The U.S. Army Group Examination Alpha
Bregman: Revision of Army Alpha Examination
Wells: Revised Alpha Examination (form 5)
Thorndike: The Thorndike Intelligence Examination for High School Graduates
Otis: S–A Tests of Mental Alertness (20 minutes)
Pressey: Senior Classification Test and Senior Verifying Test
Bureau VI: Mental Alertness

PERCENTILE	CAVD	ARMY ALPHA	BREG. ARMY	WELLS ARMY	THORNDIKE	OTIS A	OTIS B	PRES. VERIF.	PRES. CLASS	BUR. A	BUR. B
99	436.0	207.3	206.0	222.0	103.0	70.8	71.2	95.2	95.5	173.7	173.0
95	432.0	198.0	198.0	202.5	94.4	68.3	65.0	92.7	93.6	164.0	166.7
90	428.6	190.0	194.0	196.7	91.9	67.0	62.0	91.3	91.2	155.0	155.0
85	425.7	183.3	190.0	192.2	88.3	65.2	60.0	90.0	89.0	150.0	139.0
80	423.9	177.8	186.4	186.0	84.1	62.0	57.6	87.6	87.0	145.0	135.0
75	422.3	173.3	182.7	178.0	82.3	56.7	54.0	85.5	85.3	139.2	123.6
70	420.8	168.0	178.3	170.0	80.5	55.3	50.1	83.6	83.0	135.8	120.7
65	417.0	160.0	171.7	167.1	74.2	54.0	48.4	81.9	79.8	128.8	116.3
60	414.3	156.7	167.3	164.3	70.8	52.3	47.0	80.1	77.4	126.3	112.0
55	411.4	153.3	163.6	161.4	67.5	50.4	45.7	78.4	75.0	120.0	106.7
50	408.9	150.0	160.0	156.7	64.2	48.0	41.6	75.8	71.6	117.8	103.1
45	406.7	145.6	155.0	150.0	60.8	46.7	39.9	69.0	70.5	115.6	100.6
40	404.4	141.1	150.0	145.6	57.9	45.3	37.8	66.0	67.5	111.3	96.3
35	402.2	135.0	144.3	141.1	55.0	43.9	35.3	62.1	61.5	107.5	87.5
30	400.0	128.0	138.8	132.5	52.1	42.4	32.3	60.4	58.0	102.5	80.0
25	398.0	120.0	133.8	125.7	48.3	40.5	29.6	57.0	52.0	95.0	68.3
20	396.0	115.0	120.0	120.0	43.0	38.5	28.1	54.6	48.0	86.7	65.0
15	392.5	100.0	107.5	106.7	39.5	36.5	26.3	51.8	43.5	75.0	57.5
10	384.2	86.7	96.7	100.0	37.5	33.8	23.0	46.0	37.0	62.5	47.5
5	380.8	70.0	70.0	75.0	35.5	26.0	18.0	33.0	33.8	46.7	41.3
1	369.0	38.0	44.0	58.0	29.0	17.4	8.4	26.4	17.4	24.0	19.0

This assumption will be treated in a later paper. The conversion tables from test to test should also be computed from regression formulae relating such tests. The requests for this material have been so frequent, however, that the table of this paper has been presented as a preliminary set of transmutation values.

| | | |

MULTIPLE CORRELATION AND THE MULTIPLE FACTOR METHOD

HARRY PELLE HARTKEMEIER
University of Missouri

AND

LAURA MAE BROWN
Christian College

THE method of multiple correlation seems to have been largely ignored or neglected by educational statisticians in their efforts to evaluate a "battery of tests" in a psychological examination which has been used to measure or forecast scholastic ability or performance in college. Simple correlation, tetrad differences, etc., have been used, and Professor L. L. Thurstone has put forward his multiple factor method to enable one to discover the presence and extent of one or more factors in a "battery of tests." The method of multiple correlation may be used to advantage in evaluating psychological examinations. This method is similar to the multiple factor method in some respects, but differs sharply from it in several ways. It is interesting as well as instructive to apply the two techniques to the same data.

This study was made to determine (1) if there is a significant relationship between the scores made on psychological examinations given to students entering college and the grades they received in all their various courses of study, (2) if the scores made on the psychological examination were the result of one or more factors, and (3) if each part of the psychological examination was equally valuable or significant in forecasting scholarship.

The psychological examination used was prepared by Professor L. L. Thurstone and Thelma Gwinn Thurstone of the University of Chicago and published by the American Council on Education for use in 1931. The following table gives

information regarding the nature of the tests and the score possible on each test.

TEST	NAME	POSSIBLE SCORE
B	Completion	80
C	Artificial Language	73
D	Analogies	56
E	Arithmetic	80
F	Opposites	81
G	Total or Gross Score on Examination	370

The letters were assigned to the tests in order to refer to them readily; the letters will be used as symbols to represent the tests in equations and tables which follow in this study. The grades used in this study were those of sixty-one students who entered Christian College in September, 1931, and were given the test by college officials that Fall, and who completed four complete semesters of college work at Christian College in May, 1933. The grades, to be represented by the letter A, were stated as average points per hour, with 300 the maximum possible and 0 the minimum.

The coefficients of correlation between the variables are given in Table I.

TABLE I

Coefficients of Correlation Between A, Average Grades, B, C, D, E, and F, Tests in a Psychological Examination, and G, Total Score on All Tests in a Psychological Examination

	B	C	D	E	F	G
A	.4142	.4275	.4947	.4103	.5832	.6460
B		.5599	.3041	.3294	.6855	.8030
C			.2841	.3189	.6137	.7754
D				.2952	.3509	.5938
E					.2887	.5903
F						.8450

All of the coefficients of correlation in Table I are significant[1] (or reliable) and all but five (those under .328) are highly significant. This does not mean that there is a high degree of relationship between the total score on the psychological examination and the scholastic record, for the coefficient of correlation is only +.65, but it does mean that the relationship which is present in this set of data may be expected to occur in other years to a similar degree. The score on part of the examination (one test on Opposites) is almost as highly correlated (+.58) with college scholarship as is the total score. It is possible to predict college scholarship on the basis of this one part (Opposites) almost as well as it can be predicted on the basis of all five tests in the psychological examination. Thus, F seems to be the most important part of the examination. Which of the other tests is the most important? Which part adds most to the knowledge obtained from test F? Do some of the tests merely duplicate part of the information obtained in test F without adding anything to it? Is F the most important test merely because it was possible to obtain a larger score on this test than on the other

[1] The reader should recall that the word *significant* has a technical meaning when used in the field of statistics. Its technical meaning is more restricted than the meaning of the word as generally used. When applied to statistical constants the word *significant* means reliable or consistent rather than important or useful. To say that a value is significant means that it (or a larger value) could arise due to chance alone when the true value is zero only 5 (or less) times out of 100. We can expect to obtain similar values (values which are consistent with the value we have) from additional data (samples). The probability that our value is spurious or a random fluctuation from zero is .05. This probability is considered small enough to ignore. To say that a value is highly significant means that it (or a value larger) could arise due to chance alone only about 1 time in 100. The probability that our value is spurious or unreliable is .01. Such values may be trusted with confidence; you may ignore the 1 chance in 100 that your result is unreliable. For 61 observations, or 59 degrees of freedom, coefficients of correlation are significant if they equal (or exceed) .252 and are highly significant if they equal (or exceed) .328. See *"Correlation and Machine Calculation,"* by H. A. Wallace and G. W. Snedecor, p. 63.

tests? Is there any relationship between the possible score on a test and the importance of that test in the group of tests which make up the examination? These are questions which we seek to answer by applying the method of multiple correlation. What is the relationship between the dependent variable A (average grade) and the five independent variables B, C, D, E, and F (scores obtained on five tests in a psychological examination)? As a graphic analysis did not reveal much evidence of curvilinear relationship A was written as a linear function of the other variables

$$A = a + b_{AB.CDEF}\,B + b_{AC.BDEF}\,C + b_{AD.BCEF}\,D + b_{AE.BCDF}\,E + b_{AF.BCDE}F. \tag{1}$$

This equation was not fitted to the original data in this form, for it is a simpler process to fit the function in the form which results when the variables are expressed in standard units (as deviations from the arithmetic means in units of the standard deviations). Thus, if $A' = \frac{A - M_A}{\sigma_A}$, $B' = \frac{B - M_B}{\sigma_B}$ etc., the equation becomes

$$A' = \beta_{AB.CDEF}\,B' + \beta_{AC.BDEF}\,C' + \beta_{AD.BCEF}\,D' + \beta_{AE.BCDF}\,E' + \beta_{AF.BCDE}F'. \tag{2}$$

This function may be fitted to the data by solving the five normal equations simultaneously:

$$\left.\begin{aligned}
r_{AB} &= \beta_{AB.CDEF} + \beta_{AC.BDEF} r_{BC} + \beta_{AD.BCEF} r_{BD} + \beta_{AE.BCDF} r_{BE} + \beta_{AF.BCDE} r_{BF} \\
r_{AC} &= \beta_{AB.CDEF} r_{BC} + \beta_{AC.BDEF} + \beta_{AD.BCEF} r_{CD} + \beta_{AE.BCDF} r_{CE} + \beta_{AF.BCDE} r_{CF} \\
r_{AD} &= \beta_{AB.CDEF} r_{BD} + \beta_{AC.BDEF} r_{CD} + \beta_{AD.BCEF} + \beta_{AE.BCDF} r_{DE} + \beta_{AF.BCDE} r_{DF} \\
r_{AE} &= \beta_{AB.CDEF} r_{BE} + \beta_{AC.BDEF} r_{CE} + \beta_{AD.BCEF} r_{DE} + \beta_{AE.BCDF} + \beta_{AF.BCDE} r_{EF} \\
r_{AF} &= \beta_{AB.CDEF} r_{BF} + \beta_{AC.BDEF} r_{CF} + \beta_{AD.BCEF} r_{DF} + \beta_{AE.BCDF} r_{EF} + \beta_{AF.BCDE}
\end{aligned}\right\} \tag{3}$$

These normal equations involve only the coefficients of correlation, which have already been obtained. The normal

equations required for the fitting of (1) involve sums of squares and products derived from the original data, A, B, C, D, E, and F. After (2) has been fitted to the data it is easy to obtain (1) from (2). Both $b_{AB.CDEF}$ and $\beta_{AB.CDEF}$ are regression coefficients which measure the relationship between A and B when the effects of the other variables (C, D, E, and F) have been eliminated. This does not mean that the other variables are ignored. Their influences are simultaneously measured and eliminated so far as $b_{AB.CDEF}$ and $\beta_{AB.CDEF}$ are concerned. Even though they are both regression coefficients $b_{AB.CDEF}$ is given in original units while $\beta_{AB.CDEF}$ is given in abstract units. When different units of measurement are used[2] the b's change, but the β's remain the same. This means that the β can be compared with each other with respect to size while the b's should not be compared.

After (2) has been fitted to the data, it becomes

$$A' = -.0711(\pm .1400)\ B' + .0491(\pm .1290)\ C' + .2866(\pm .1071)\ D' + .2060(\pm .1071)\ E' + .4417(\pm .1471)\ F' \qquad (4)$$

The coefficient of multiple correlation between A on the one hand and B, C, D, E, and F on the other hand is $R_{A.BCDEF} = .6896$. This value is highly significant or reliable[3] even though it does not indicate a high degree of relationship. The values given in parentheses are the standard errors of the β's, only two of which are more than twice their standard errors. Thus F and D are reliably associated with scholastic records. The coefficients of simple correlation between the other three variables and scholarship were about the same size, but the β's connecting B and C with A are small and are not signifi-

[2] Such as inches instead of feet, cents instead of dollars, tons instead of pounds, 1000 bushel units instead of bushels, etc. In this study if the A values were all multiplied by 10 so that the grades varied from 0 to 3000 instead of 0 to 300, all the b's would change but the β's would not be different. Thus the β's are free from the effect of different amounts of variation in the different variables included in the equation.

[3] See Wallace and Snedecor, *op. cit.*, p. 63, Table 16. In this case there are 6 variables and 55 degrees of freedom.

cantly different from zero.[4] The beta measuring the relationship between E and A is 1.92 times its standard error. This is not quite large enough to be accepted as proof of a significant and reliable relationship, but it is very close to the value (2.004σ) that would be accepted.

Since two of the variables clearly have no significant association with scholarship the function was refitted excluding these two variables (B and C) to obtain

$$A' = +.2858(\pm .1053)D' + .2035(\pm .1030)E' + .4241(\pm .1050)F' \quad (5)$$

The coefficient of multiple correlation (.6872) is almost as large as it was when B′ and C′ were included. The coefficient of E′ is still not quite large enough in terms of its standard error to be judged clearly significant, for it is 1.976 times its standard error; however, it is so close to being significant that it may well be considered.

Now, what do these values mean? The scholastic records of the students varied and we wish to explain this variation and predict it if possible on the basis of the examination scores. The variation or tendency of the scholastic records (A) to differ may be measured by V_A (termed variance and is equal to σ_A^2). We would like to explain all of the variance V_A. What part can we explain in this particular case? The answer to this question is given by $R^2_{A.DEF} = .4722$. We have explained 47 per cent of the variance in scholastic performance or record in terms of the scores obtained by the students in tests D, E, and F. Tests B and C had no reliable relationship with scolastic records. This does not mean that the students did poorly on tests B and C, for on the contrary, they did well on these tests. On test C the students as a whole made a higher score than they did on any other test. This was not because the possible score on test C was higher than the possible score on the other tests, for three other tests had

[4] It would have been difficult to explain the negative sign before $\beta_{AB.CDEF}$ if this value had proved to be significant. However, as the standard error is larger than the beta, the latter value may be dismissed as a random fluctuation from zero.

a larger possible score than test C. It happened that several good students obtained low scores and several poor students did well on tests B and C. This means that these tests were not so satisfactorily devised to reveal or forecast good scholarship in college as were the other tests.

Of the three test grades which were associated with good scholarship, which test was the best one for predicting scholarship? From the size of the betas in (5) it appears that F is the most valuable test in forecasting scholarship. D is next in importance with E last of the three tests. However, one must guard against the tendency to conclude that F is twice as important as E since $\beta_{AF.DE}$ is about twice the size of $\beta_{AE.DF}$. This is a controversial statistical question into which it is not advisable to enter in this paper, but since the betas do not add up to anything there is some reason for hesitating to attach precise meaning to their relative size. However, it is possible to obtain coefficients of determination which may be used to answer the question of the relative importance of the tests. These values are obtained by multiplying each beta by the corresponding coefficient of correlation; for example, $\beta_{AF.DE}r_{AF}$ may be used as a measure of the relative importance of F in the explanation or forecasting of scholarship. This use arises out of the fact that the βr's add up to R^2, and R^2 measures the percentage of the variance in scholarship which we are able to explain or forecast on the basis of the psychological examination.

$$\left.\begin{array}{l} R^2_{A.DEF} = \beta_{AD.EF}r_{AD} + \beta_{AE.DF}r_{AE} + \beta_{AF.DE}r_{AF} \\ .4722 = .1414 + .0835 + .2473 \end{array}\right\} \qquad (6)$$

On this basis F is almost three times as important (or valuable) as E in predicting scholarship. We are able to explain only 47.22 per cent of the variance in scholarship using the scores on three tests, and the scores on F account for (enable us to forecast) 24.73 per cent of the variance in scholarship. Before the betas may be compared to determine the relative importance of the tests, the betas should be

squared,[5] just as coefficients of correlation must be squared before they can be properly interpreted as a percentage or proportion.

The multiple regression equation in original units—in form (1) rather than (2) is

$$A = 42.8381 + 1.4843D + 1.0455E + 1.5673F \qquad (7)$$

These regression coefficients, which are in original units, are quite different from those in (5), which are in standard units. The coefficients in (7) should not be compared with each other, for even though they are (in this case) all in the same units (scores on tests) the variation in F is much larger than the variation in D and E. However, equation (7) is useful because we can estimate scholarship by inserting the psychological test scores directly in place of D, E, and F to obtain A, while we could not substitute the scores in (5) without subtracting the average of each test for all students from the scores obtained by one student and dividing these differences by the corresponding standard deviation of all scores on each test. Even if this were done the answer, A′, would not be the scholastic rating but a deviation from the average of the scholastic rating of all students and expressed as a multiple of the standard deviation of all scholastic ratings. It is more convenient to use equation (7) to estimate scholastic performance from test scores.

Although the coefficient of multiple correlation was highly significant or reliable, it was not large enough to indicate good correlation between scores on the psychological examination and scholastic performance. $R_{A.DEF} = .6872$. This indicates that only 47 per cent of the variance in scholarship can be explained in terms of scores on the psychological examination. This small percentage is probably due to several factors or special circumstances, as follows:

1. The grades in each course are rough measures of scholarship, for only nine different grades were given—300,

[5] See H. P. Hartkemeier, "*The Supply Function for Agricultural Commodities*," University of Missouri Studies, Vol. VII, No. 4, Appendix II.

275, 225, 200, 175, 125, 100, 75, and 0. Of course A shows more variation than these nine values, for A is an average of a number of grades for each student. The series of values measuring scholarship varies from 12 to 299 by small intervals (varying in size).

2. The relationship between grades and ability is probably an S curve rather than a straight line. The correlation technique used above assumes a straight line relationship.

3. The students whose work is not satisfactory are placed upon the "unsatisfactory list," have little to do but study, and receive extra attention.

4. Some of the grades used in determining A were received in courses begun 1½ years after the test was taken. The scholastic record of the first year could probably have been predicted with a higher degree of accuracy than the scholastic record over two years.

The fallacy in considering possible scores or actual scores on tests and simple coefficients of correlation between test scores and scholastic records to determine the relative value of tests will be apparent after a study of Table II.

TABLE II

Summary of Information Regarding Tests Scores and Scholastic Records

1	2	3	4	5	6	7
Test	Possible Sc. as part of total	Actual Sc. as part of total	Simple Cor. with A, Scholarship	Regression Coefficients with A		Coefficients of Determination βr
				Orig. Units b	Abstr. Units β	
B	.2162	.1834	.4142	Not significantly different from zero		
C	.1973	.2730	.4275			
D	.1514	.1960	.4947	1.4843	.2858	.1414
E	.2162	.1274	.4103	1.0455	.2035	.0835
F	.2189	.2202	.5832	1.5673	.4241	.2473
	1.0000	1.0000				

Although test C was fourth in the order of possible points the average score on it was higher than the average scores on the other four tests. The simple coefficient of correlation between test C scores and scholastic records is highly significant or reliable; nevertheless, test C adds practically nothing to the knowledge obtainable from tests F, D, and E. Similarly, the correlation between test B scores and scholastic records is highly significant, but test B adds very little to the other tests so far as predicting scholarship is concerned. Tests B and C are more highly correlated with F than with D and E so we may conclude that B and C have more in common with (measure the same ability as) F than with D and E. Even after the correlation of the other tests with scholastic work has been allowed for, tests D and E are still significantly correlated with scholastic record. This means that tests D and E each measure abilities which are important for good scholarship and which are not measured by tests B, C, and F. If we try to determine the relative importance of tests F, D, and E on the basis of simple coefficients of correlation as shown in column 4 we would be decidedly misled as is apparent from column 7.

The multiple factor method leads to the same conclusion on some questions raised regarding the psychological examination. This method was devised to determine if more than one factor "must be postulated in order to account for the intercorrelations"[6] between the tests, and to measure the extent to which each factor was present in each test by determining "factor loadings."

Space does not permit a detailed explanation of this method, but the reader may find it in the references given. The computations were made according to the outline given in "A Simplified Multiple Factor Method,"[7] by L. L. Thurstone, and the factor loadings are given in Table III.

[6] L. L. Thurstone, "*The Theory of Multiple Factors*," p. iii.

[7] Distributed by the University of Chicago Book Store.

TABLE III

Factor Loadings Determined by the Simplified Method (Center of Gravity Method)

TEST	FACTORS					
	I	II	III	IV	V	h^2
B	+.7830	−.2045	+.0890	+.1864	+.0661	.7019
C	+.7299	−.1571	−.2098	−.1191	+.0672	.6202
D	+.4840	+.2726	+.1317	−.1238	+.1254	.3569
E	+.4768	+.2760	−.0796	+.0791	+.0155	.3163
F	+.8013	−.2644	+.1603	−.1179	+.0818	.7583
Σk^2	2.2495	.2869	.1013	.08441	.03154	
$\frac{\Sigma k^2}{n}$	.4499	.05738	.02026	.016882	.006308	

The meaning of the factor loadings may be briefly indicated by several paragraphs from "The Theory of Multiple Factors,"[8] by L. L. Thurstone:

"In all of our present work on factor theory we shall assume that the factors combine in a linear manner in order to produce the total performance in a test. This is the assumption of the usual regression equation. The general pattern that we shall follow in factor analysis can be illustrated by a linear equation for two general factors as follows:

$$S_a = a_1 x_1 + a_2 x_2 \qquad (1)$$

in which

S_a = standard score in test A.
x_1 = standard score in ability no. 1.
x_2 = standard score in ability no. 2.
a_1 and a_2 are coefficients.

The two abilities are assumed to be uncorrelated. The performance S_a, on a test, is here expressed as a function of the scores in the two general abilities. If the value of a_1 is large

[8] Photo-lithoprint reproduction of author's manuscript by Edwards Bros.

and the value of a_2 is small, then we should interpret the equation by saying that test *A* calls for much of ability number 1 but very little of ability number 2.

"Inspection of equation (1) shows that a good performance S_a, on a test, may be accounted for by superiority in either of the two abilities that the test demands, or else by fair ratings on both of the abilities. The same reasoning may be extended, of course, to the situation in which a test performance is determined by three, four, or more abilities. Hence, a good performance on a test does not prove that the subject has superior rating on any one ability except in the case in which a test is assumed to represent only one ability. The latter is the special case to which Spearman and his students have given much attention.

"The coefficients a_1 and a_2 are essentially properties of the test in that these coefficients indicate the extent to which test *A* calls for each of the abilities 1 and 2. We shall see that these coefficients have certain very definite restrictions in possible magnitude. The values of x_1 and x_2 are attributes of the individual subject who makes the performance S_a.

"We shall consider the combination of equation (1) which refers to test *A* with a second similar equation for another test *B*. We have then

$$S_a = a_1 x_1 + a_2 x_2$$
$$S_b = b_1 x_1 + b_2 x_2$$

in which: S_b = standard score in test *B*. b_1 and b_2 are the factor loadings for test *B* that correspond to a_1 and a_2 for test *A*. . . . . ."

In this case five factors have been used to explain the intercorrelations of scores on five tests, so five equations are necessary. The method by which these equations are used to determine the factor loadings is called a least square solution, but this does not mean that the five equations are solved simultaneously as in the method of multiple correlation. If this were the case it would be necessary to decide at the beginning how many factors are present in order to set up the equations and determine the factor loadings. Furthermore, if the number of factors were increased or decreased it would become necessary to resolve for the factor loadings.

After discussing such a least square solution Professor Thurstone writes:[9]

> "The least square solution that we have considered so far is illuminating in some respects but it is tedious and in fact prohibitive in labor when the problem involves more than six or eight variables. We shall now consider another form of the least square solution which is practically useful no matter how many tests or factors may be involved. Our present solution is in general as follows. We shall determine one factor at a time. We ask first what the factor loadings must be, for one factor only, so that the discrepancies will be minimized. . . . When these first factor loadings have been found, we can be certain that no one can find a better linear single factor description of the given correlations. The residuals are then treated in the same manner as the given coefficients in order to find the best second factor loadings. We now find the second factor loadings which will minimize the first factor residuals."

Professor Thurstone later devised a further simplified method[10] and published it in "A Simplified Multiple Factor Method." The factor loadings given in Table III were determined by this last method. If only three factors are deemed necessary the loadings for factors IV and V may be omitted without changing the loadings for factors I, II, and III. No mathematical or objective method is given by Professor Thurstone for determining when enough factors have been determined. He even discourages the use of such methods, for he writes:[11]

[9] L. L. Thurstone, *op. cit.*, p. 28.

[10] These processes are analogous to the graphic method in multiple correlation in contrast to the mathematical method. More exactly, perhaps, they are analogous to the method of part correlation in contrast to the method of multiple correlation. The authors do not wish the reader to conclude that because they have used the simplified method they approve the graphic method in multiple correlation or the method of part correlation. They do not approve of them except as preliminary steps. Neither should the reader conclude that the authors approve of the simplified multiple factor method. The latter is used because Professor L. L. Thurstone approves of it.

[11] L. L. Thurstone, *op. cit.*, p. 27.

> "The computation of probable errors is not likely to be very profitable. It is probably better to calculate the correlation coefficients in terms of the rotated factors, to reduce their number as far as possible, and then to compare the calculated coefficients with the original correlation coefficients. If the number of factors is small, compared with the number of tests or variables, and if the discrepancies between the calculated and observed coefficients are of an order of magnitude comparable with the probable errors of the given correlations, then the factorial description of the tests is satisfactory and convincing even tho we do not trouble to compute probable errors of the factor loadings themselves. It will probably always happen that a set of tests contains some common factors of small significance which can be ignored and which are of course not accounted for by the probable error formulae."

Without probable errors it is difficult, if not impossible, to determine whether or not a set of factor loadings is significantly or reliably related to scholastic performance or ability. Professor Thurstone tells us to reduce the number of factors "as far as possible." How far is this? Without probable errors it is difficult to determine whether to add another factor or drop one already included. When Professor Thurstone states, "It will probably always happen that a set of tests contains some common factors of small significance which can be ignored and which are of course not accounted for by the probable error formulae," he seems to be confused on the subject of probable error and reliability of statistical (or quantitative) values determined by the use of samples. Does *of small significance* mean small in size or small in relation to a measure of reliability? If the latter is meant he is right, but if the former is what he intended the words to convey he should be reminded that frequently a small statistical constant is more reliable than a large one. This occurs when the small value is a larger multiple of its standard error than the large value (is a multiple of its own standard error). Thus, the small factors present may be very reliable and can be depended upon to recur in additional samples while the

large factors which are not twice their standard errors (3 times their probable errors) may not be expected to recur in additional samples. Returning to the correct meaning of the words *of small significance,* the authors wish to point out that even if Professor Thurstone did mean small in relation to a measure of reliability (the only logical meaning) it is difficult to determine what he meant by the rest of the sentence, " . . . and which are of course not accounted for by the probable error formulae." If some factors present are of small significance because they are small in relation to their probable errors, they may be ignored *because* they may be easily accounted for as random fluctuations which are measured (accounted for) "by the probable error formulae"! The purpose of having a measure such as the probable error is to be able to determine the extent of random fluctuations—those deviations present in the sample which cannot be regarded as significant or reliable.

In the last quotation from Professor Thurstone's work some help seems to be given in the problem of determining how many factors to include or determine by the statement, " . . . if the discrepancies between the calculated and observed coefficients are of an order of magnitude comparable with the probable errors of the given correlations, then the factorial description of the tests is satisfactory and convincing . . . " In order to see if the distribution of the discrepancies enables one to determine how many important or significant factors are present Table IV was compiled.

It is clearly evident that as more factors are added the distribution of errors narrows down considerably until four factors are evaluated. The averages given at the bottom of Table IV were determined from the actual values of the discrepancies and not from the frequency distributions, for with only 10 observations it is not advisable to compute the averages from the frequency distributions, which are used merely to show in general how the variation in the discrepancies decreases. The question is, "At what point are the dis-

TABLE IV

Frequency Distributions of Discrepancies Between Observed Coefficients of Correlation and Those Computed by Using 1 to 5 Factors (or Sets of Factor Loadings)

CLASS INTERVALS	1 Factor	2 Factors	3 Factors	4 Factors	5 Factors
−.0975→* −.0825	1				
−.0825→ −.0675	2				
−.0675→ −.0525					
−.0525→ −.0375	1	1			
−.0375→ −.0225	2	1	1		
−.0225→ −.0075	1	4	3	3	2
−.0075→ +.0075		1	3	6	7
+.0075→ +.0225		2	2	1	1
+.0225→ +.0375	1	1	1		
+.0375→ +.0525					
+.0525→ +.0675	2				
+.0675→ +.0825					
+.0825→ +.0975					
Average discrepancy:					
Regarding signs	−.02076	−.00593	−.00384	−.00214	−.00141
Disregarding signs	.05102	.01991	.01324	.00662	.00547

* Means "up to but not including."

crepancies small enough to indicate that the factorial description is satisfactory?" The probable errors of the largest, average, and smallest observed coefficient of correlation are .0678, .1072, and .1177, respectively. From these probable errors it would seem as if only one factor is present since only a few of the discrepancies are larger than the first probable error given and all of the discrepancies are smaller than the last two probable errors. However the factor loadings given in Table III for the second factor are fairly large and the use of two factors causes an appreciable reduction in the discrepancies, as shown in Table IV. It seems as if the probable errors of the coefficients of correlation will not be of much help in evaluating the factor loadings or testing their significance. The conclusion appears to be that Professor Thurstone has not given a means of telling when to stop in the

determination of additional factors after the first. Indeed he states that " . . . additional factors can be added as long as one has patience and interest."[12] The authors believe that something besides patience and interest are necessary to establish the presence of a factor quantitatively. All statistical constants should be evaluated in terms of their probable errors. The authors believe that at least a mild attack of what Professor Thurstone refers to as the "probable error phobia"[13] is beneficial to all quantitative workers (including Professor Thurstone).

Professor Thurstone states that, "The significance of any factor may be estimated by the summation $\Sigma k^2/n$ which becomes smaller and smaller." The values referred to are given in Table III. The values for the second and third factors seem small but the values for $\Sigma k^2/n$ which Professor Thurstone gives in "A Simplified Multiple Factor Method" are .525, .066, and .035, so the second and third values are small in his illustration as well as in ours. In Table III the values seem to indicate the presence of two factors, but there is some doubt about the third factor. There is not sufficient evidence of 4 or 5 factors being involved in the examination scores.

Does the multiple factor method enable one to determine which tests are the most important in the prediction of scholarship? This point is not discussed at length by Professor Thurstone, but it appears to the authors that h^2 is meant to give some evidence on this subject. The h^2 value for any test is obtained by adding the squares of the factor loadings for that test. The first value for h^2, .7019, is obtained by adding $(.7830)^2$, $(.2045)^2$, $(.0890)^2$, $(.1864)^2$, and $(.0661)^2$. Professor Thurstone terms h^2 the communality of the test and states,[14] "The communality for each test is the sum of the squares of its factor loadings. The fact that h^2 is less than unity means that the test has either a specific factor unique in the group of seven tests [his illustration included seven

[12] *Ibid.*, p. 44.
[13] *Ibid.*, p. v.
[14] *Ibid.*, p. 22.

tests while ours has five tests] or else has a chance error factor in its score, or both of these factors." The h^2 values given in Table III rank the tests in the order F, B, C, D, E. The same order or rank of the tests is obtained from the h^2 values when one, two, three, four, or five factors are evaluated. The h^2 values change, of course, when the number of factors changes from one to five, but the five sets of h^2 values all rank the tests in the same order. The method of multiple correlation ranks the tests in a different order, placing D and E before B and C—the last two being not significantly different from zero in importance for predicting scholarship. Here we see a decided difference in the two methods. The method of multiple correlation measures the specific contribution of each test to our knowledge of scholarship and distributes all joint contributions (common factors) equally to the tests involved, while the multiple factor method measures to what extent the common factors are present in each test, ignoring the specific contribution of each test. Furthermore, in the case of common factors, it seems as if the multiple factor method gives each test a loading which measures the extent to which the common factor is present instead of distributing equally to the tests anything which is common to them. In other words, if three tests have something in common so far as predicting scholarship is concerned the multiple factor method gives each test a value based upon this factor, while the method of multiple correlation splits up the joint action into three parts, giving each test its share in the joint influence or factor. The multiple factor method lumps the specific factors and random fluctuations (chance errors) together as something not directly important and separates in the measuring or evaluating process the common factors determined by factor loadings. The method of multiple correlation combines or lumps the specific factors and a part of all common factors into a measure (regression coefficient, etc.) for each test leaving out only the random fluctuations (chance errors).

The method of multiple correlation seems superior in that it measures the specific contribution of each test as well as the joint contributions of all tests, while the multiple factor method seems superior in that it measures the presence of each common factor separately from other common factors and distributes the common contributions of tests unequally among the tests—presumably in *proper* proportion (although how this may be measured or judged is not clear). However, the multiple factor method does not give us a basis for judging how well a battery of tests (individually and as a group) enables us to predict scholastic success in college while the method of multiple correlation does give us such a basis. So far as the authors can determine the multiple factor method does not give any means of judging whether or not a "battery of tests" is a good one for forecasting scholarship.

The factor loadings are evaluated in terms of our ability to reproduce (or compute) the inter-correlations of test scores by using the factor loadings while the betas are evaluated in terms of our ability to estimate scholarship from the test scores. We have only one beta for each test, but we have as many factor loadings for each test as there are common factors in the test scores. Thus the betas may be thought of as a combination of the three or five sets of factor loadings into one set if we add to the combination the abilities which are measured by only one test (specific factors) and make the combination in such a way that anything common to two or more tests is distributed equally among the tests involved.

It is necessary to have scholastic records of the students who took the psychological examination in order to apply the method of multiple correlation while the scholastic records are not necessary for the application of the multiple factor method. This might be thought of as an advantage of the latter method in so far as it is not necessary to wait two years (in this case) to learn if the examination tests were good ones, but it must be pointed out that the purpose of the examination is to forecast scholastic ability and as long as you have not compared the

forecast with the performance you do not know how good the forecast is. Professor Thurstone points out:[15]

> "There is a rather general preference for a set of tests in a battery to contain only one common factor. A battery of tests is criticised when it is found that there are group factors present or that more than one general factor is needed to account for the intercorrelations. This idea is defensible in some situations and not in others. *If the investigator is assembling a set of tests for the prediction of college scholarship, then the criterion of success is of course the prediction of scholarship. If he succeeds in doing that, then the battery of tests is a good one, quite irrespective of whether it requires more than one factor to describe the intercorrelations.* (Italics ours.) There is nothing intrinsically good or bad about the presence of more than one common factor. If, on the other hand, the investigator is trying to assemble a battery of tests that shall measure a single mental trait, then he prefers to have them so constructed that one general common factor shall be conspicuous. But group factors may still be present without doing any harm whatever."

The multiple factor method tells you merely whether or not you have constructed the tests successfully *the way you wanted to construct them*—with only one common factor (or two or more common factors) present. The method does not tell you whether the tests are good or bad so far as predicting scholarship is concerned. Answering the last question is the outstanding superiority of the method of multiple correlation.

A final point is worth making. The method of multiple correlation indicates whether the weighting assigned each test is too large or too small. In this case test D (Analogies) was given the smallest weight, whereas it should have been given the second largest weight in the determination of scholastic ability and tests B and C should have been omitted from the examination entirely (given zero weight instead of 80 and 73 respectively). The multiple factor method seems to give no help on this point.

[15] *Ibid.*, p. iv.

PRELIMINARY RESULTS OF A STUDY OF MOTHERS' OPINIONS OF CHILDREN'S RADIO PROGRAMS

H. P. LONGSTAFF
University of Minnesota

IN an earlier investigation concerning the effectiveness of children's radio programs, it was found that these programs were very popular with children and apparently quite effective. The question arises, what do mothers think of these programs? The purpose of this study was to investigate further the effectiveness of children's radio programs, to discover, if possible, any trends of opinion which may be held by mothers concerning these programs, and to determine the advisability of conducting a survey involving a larger sample of the population.

One hundred ninety-four subjects were used in this study. The information was obtained by the personal interview method. The interviewers were seniors and graduate students specializing in advertising at the University of Minnesota. It is doubtful that 194 mothers are an adequate sample of the total population, yet certain trends disclosed by this study are sufficiently definite to warrant consideration and to indicate the desirability of securing a larger *sample*. This larger sampling is being made at the present time and the data should be available shortly.

The results for the first part of this study dealing with the effectiveness of this medium of advertising are presented in the following table which is practically self-explanatory. It is apparent that the programs are very effective in inducing children to ask their parents to purchase the commodities advertised by these programs. This is in agreement with results

[1] Longstaff, H. P. Effectiveness of Children's Radio Programs. J. of Applied Psychology. April, 1936.

TABLE I

	ASKED TO BUY	BOUGHT	BOUGHT JUST TO GET THE PREMIUM	LIKED THE COMMODITY	CONTINUED TO BUY THE COMMODITY
	per cent	*per cent*	*per cent*	*per cent*	*per cent*
Coco-Malt	51	88	47	88	46
Skelly Gasoline	31	66	80	57	26
Wheaties	65	94	56	80	80

This table is read as follows—51 per cent of the mothers interviewed had been asked by their children to buy Coco-Malt. Of this 51 per cent, 88 per cent bought it, 47 per cent bought it only to get the premium, 88 per cent liked the commodity, and 46 per cent continued to buy it.

reported by the writer in a previous study. In addition to persuading the children to ask their parents to buy the commodity, the children were so insistent that a large per cent of the parents did buy them.

It is of interest to note that one of the major interests of the child in having his parents buy the commodity was not the commodity itself but the premium the child could get by sending in a box top or something of the sort. This may explain in part the rather large number of customers who bought some of the commodities once but did not buy them again. In the case of Coco-Malt and Skelly Gasoline, only about half of those who bought these commodities as a result of the children's requests are now using the commodity. Wheaties, on the other hand, shows little of such loss.

The reasons why parents continue to buy one product and not another cannot be determined from these data, although there seems to be some basis for the conclusion that the nature of the product is probably quite influencial in this connection. Wheaties is one of a group of staple food products, consequently once purchased there is little reason to change to another brand. Coco-Malt, on the other hand, is more of a luxury and can easily be dispensed with. Skelly Gasoline probably is little better or little worse than any other gasoline

on the market and since it is more difficult to obtain due to the relatively small number of stations owned by this company as compared to Standard Oil or some of the other more popular brands, this factor of convenience may well be the determining factor with regard to its inability to hold customers. Nevertheless, the fact that both Coco-Malt and Skelly have held roughly fifty per cent of the customers obtained as a result of the children's radio programs is significant.

It is hoped that the larger study mentioned above and now in progress may afford more specific answers to this question.

The results of the second part of the present study also show that children's programs are very popular. Eighty-six per cent of the mothers interviewed indicated that their children listened to these programs, 5 per cent said they did not, and 9 per cent did not know.

The mothers were then asked if these programs produced any bad effects in their children: 34 per cent answered yes; 22 per cent indicated the programs made their children nervous; 8 per cent said they caused disturbance of the child's digestion; 9 per cent said they caused bad dreams; 11 per cent said they increased fear of the dark; and 44 per cent said these programs were unsuitable for children to listen to. This last is of considerable importance. Even though the programs are popular among the children, it is the parents who buy the products; consequently, it would seem that the policy of sending programs into the home which antagonize 44 per cent of the mothers is likely in the long run to give rise to a background of unpleasant feeling tone which may seriously influence future sales by these companies.

It was further discovered that 25 per cent of these mothers had attended meetings where these programs had been discussed. Foremost among such groups were Parent Teachers Associations. Eighty-eight per cent of the mothers who had attended such meetings claimed these groups considered the children's programs harmful; 8 per cent neither harmful nor beneficial; and 4 per cent were undecided; and not

one mother reported these group discussions as being favorable toward the programs. Furthermore, 32 per cent of the mothers reported that these groups had decided to take some action to improve the programs. The most frequently suggested action was boycott of the advertised products.

The facts seem pretty obvious. Any selling or advertising device which arouses such organized opposition to it is of questionable advertising and selling value to say the least.

What is it about these programs that mothers object to? The following are the criticisms leveled at the programs: children copy bad language used in program; play is more beneficial; get on parents' nerves; shooting people gives rise to bad ideas; too exciting; not good for sickly children; valueless—teaches nothing; teaches bad grammar; interferes with study; teaches bad manners; home furnishes better amusement than radio; crime called to attention too much; couldn't even talk during meals; interferes with meals; too exaggerated; too noisy; makes children too sophisticated; and children quarrel about which program to listen to.

In spite of the fact that this and the previous study reported by the writer indicate the decided effectiveness of these radio programs in selling the commodities they advertise, the writer cannot help but feel that the resentment they are arousing in the minds of some mothers may in the long run lead to very serious consequences for those companies sponsoring them. Since the number of subjects sampled in this study is small, the results may not be typical but they do indicate the importance of making the larger studies now in progress.

NEWS AND NOTES

Reports from the Federal Office of Education show that approximately 9,500 persons were rehabilitated last year under the terms of the national vocational rehabilitation act. Of this number 3,600 were given training in institutions to fit them for specific occupations. Almost 2,000 were prepared for employment of some kind by being fitted with artificial appliances such as hands, arms, and legs. Through funds provided by the Federal Emergency Relief Administration maintenance or living expenses were provided for 829 disabled persons while they were being trained for employment or undergoing physical restoration. In addition, several thousand were provided with employment consistent with their abilities and handicaps. Because of the different types of services necessary in individual cases it is impossible to rehabilitate disabled persons on a mass basis. Physical restoration is different for different cases and the training for specific employment may involve a protracted period of time.

The Second Conference on Education and the Exceptional Child was held under the auspices of the Child Research Clinic of the Woods Schools at Langhorne, Pa., on April 23, 1936. One of the principal speakers was Dr. Elton Mayo, Professor of Industrial Research, Harvard University, who declared that the difficult child is socially maladjusted, and that it it not easy to develop from happy childhood to undamaged maturity under the conditions of social instability and change existing today. He pointed out how much psychology and psychiatry had contributed to knowledge of the developing child and how important social inquiry is to the educator who attempts to utilize the findings of psychiatry, in actual handling of children. Among other outstanding authorities taking part in the Conference were David Segal, Consultant, Division of Research and Investigation, United States Office of Education; Dr. Cecil White Flemming, Director of Individual Development and Guidance, Horace Mann School, and Mrs. Rose H. Alschuler, Director, Winnetka Public School Nurseries, Winnetka, Illinois.

Reports from State departments of education, according to the Federal Office of Education, show that the enrollment in vocational education classes has, with the exception of slight decreases in the two depression years, 1933 and 1934, steadily increased since the vocational education program was organized on a national basis under the Smith-Hughes Act of 1917. The enrollment in that year was approximately 164,000. The

enrollment in 1935 in the 48 States, the Territory of Hawaii, and Island of Puerto Rico totaled in round numbers 1,248,000 youths and adults. Enrollment in various types of classes providing training for workers in the field of the trade and industries reflects to a considerable degree the trends in employment opportunities during 1935. Workers who have felt the need of acquiring additional information and skill in their respective occupations have enrolled in increased numbers in evening classes. Young persons, formerly unable to find employment because adults were given preference in available vacancies, have during the past year been employed in larger numbers. Many young persons, optimistic over the outlook for employment, have continued in school for definite occupational training in anticipation of future jobs.

Preliminary announcements have been received of the Congrès International de la Population to be held in Paris from July 28 to 31, 1937. The French Committee for the scientific study of the problems of the population is composed of the following: President, A. Landry; Vice-President, P. Rivet; Secretary, J. Alquier; Members, E. Borel, C. Bougle, J. Bourdon and F. Boverat, and others. Further information may be obtained by writing to the following address: 16, Rue de l'Estrapade, Paris (5e), France.

Dr. L. E. Beck, of the University of Oregon, has accepted an exchange appointment to the Department of Psychology, Ohio University, Athens, Ohio, with James P. Porter, for the summer session of 1936. Dr. Porter will assume his duties in the Portland Branch of the University of Oregon on June 15.

BOOK REVIEWS

SHAFFER, L. F. *The Psychology of Adjustment.* New York: Houghton Mifflin Co., 1936. Pp. 600.

FRY, C. C., and HAGGARD, H. W. *The Anatomy of Personality.* New York: Harpers, 1936. Pp. 357.

These two current publications reveal a sharp contrast between the psychological and the psychiatric approaches to personality. While it is not fair to expect the same standard of technical exactitude (since Shaffer is writing a text and Fry and Haggard present a popular discussion), the whole point of view of the two books is sharply different. Fry and Haggard drew their inspiration from psychiatric and typological approaches. Their emphasis leans heavily upon heredity. On p. 124 they comment that "Son like father, like grandfather, that is the way inheritance is transmitted; that is the way intelligence is passed." This seems to connote unfamiliarity with experimental studies of the inheritance of intelligence. Again on p. 211 we find the remarkable assertion that "No man can ever change the nature of his temperamental endowment. If he is born vivacious or phlegmatic, cheerful or irritable, anxious or moody, so he will remain." This statement seems to me to be not only an absolute contradiction of fact, but a dangerous idea to be promulgated among the general public. After years of patient work by mental hygienists, spreading the view that personality traits can be changed by proper training, to have such a view broadcast is disheartening.

Shaffer's position in this regard is much more closely in line with available data. On p. 359 he remarks, "There is no doubt of the influence of heredity on the structure of the individual, although various environmental factors modify even quite anatomical traits within certain limits. Mental ability . . . is somewhat fixed, being only moderately modifiable by cultural environment. There is reason to believe, however, that the functional characteristics of personality and adjustment are the most plastic of human traits."

It would be instructive to compare the two volumes with reference to other theoretical differences. However, we can only say that in general, Shaffer takes the view that "personality traits develop by a gradual process of learning" and interprets cases in the light of this view, while Fry and Haggard lean to the typological assumption and cast much cold water on the rôle of conflict, trial and error, and adjustment, in the determination of personal characteristics. Fry and Haggard present an entertainingly written volume, with a good selection of cases whether one agrees with the interpretation or not. Shaffer's volume I consider

a great improvement over previous texts on the normal personality, as it ties up much more closely with the concepts of general psychology. It is highly recommended for the second or third course in the customary psychology curriculum.

ROSS STAGNER,
University of Akron.

MYRTLE B. MCGRAW. *Growth. A Study of Johnny and Jimmy.* New York: D. Appleton-Century Company, 1935. Pp. xxi + 319.

Dr. McGraw's work on the twin boys, Johnny and Jimmy, has been so well publicized in both the scientific and popular literature that it is entirely unnecessary to redescribe it here. This report of her observation and experimentation with these children over a period of some two years will undoubtedly become one of the important books in the field of genetic psychology.

The author makes a division of behavior patterns into the primitive, which, according to her, function at an infracortical level and which she calls phylogenetic activities, and the more complex behavior functioning at a cortical level, which she calls ontogenetic activities. The first of these are usually evident in the neonatal period and are not easily modified by practice. The ontogenetic activities may be established through early practice at surprisingly young ages. The development or modification through practice at early ages of these ontogenetic patterns requires relatively longer time than the establishment of the same patterns at older ages. Such results might have been expected from Gesell's earlier work with the method of co-twin control. Dr. McGraw is quite certain that "wherever there is development or growth there is also a period of inception, incubation, consummation, and decline." No behavior appears fully developed, and none suddenly disappears. In the case of early appearing reflexes, such as grasping, the inception and incubation stages have probably taken place prenatally. In proceeding through this sequence of stages Dr. McGraw feels that a pattern of activity may become exaggerated, and then the excess movements are eliminated. Because of this she proposes that the contrasting theories of development which hold that behavior patterns grow from specific units, and that behavior patterns result from the elimination of excess activity, may both be correct. They are based on observation at different stages of the developmental sequence.

Most of the work reported in this book is based upon observation of the pair of twin boys although other babies have been observed. It is to be regretted that Dr. McGraw has taken the rather cumbersome verbal method of presenting her results. For many people a more liberal use of tables and charts would have made the material somewhat clearer.

There will be some also who will not subscribe to the author's theoretical interpretation. Nevertheless, this report contains valuable experimental material which other workers can use to advantage.

C. M. Louttit,
Indiana University.

Margaretta Williamson. *The Social Worker in the Prevention and Theatment of Delinquency.* New York: Columbia University Press, 1935. xvii + 236 p. $2.50.

For a number of years the American Association of Social Workers has published job analyses of special types of social work. Earlier volumes describe the duties of social workers in the family, medical and psychiatric fields, in group work, and in child care and protection. A fifth volume on vocational guidance is about to be published.

The fourth volume of the series is the one under consideration. It is divided into two parts, the first dealing with treatment by means of probation and parole, the second part deals with prevention. The work and duties of probation and parole officers, supervisors and executives are described in detail with illustrations selected from reports on such agencies in various cities. More important than treatment of delinquency is its prevention. Protective agencies, such as the Girls' Protective League of Detroit, the Big Brother and Big Sister services, and policewomen are offered as preventive measures. The functions of various workers in these agencies are described.

The value of such a book is great, although the appeal is limited: It should be of the greatest service to college guidance and placement offices, or to individuals who are interested in social work but who are unacquainted with the duties of specific positions. It should also have a decided value to workers engaged in probation, parole, and preventive work as a means of evaluating their own positions.

C. M. Louttit,
Indiana University.

Emory S. Bogardus. *Leaders and Leadership.* New York: D. Appleton-Century Co., 1934. Pp. viii + 325.

The book is evidently a text for a college course in Sociology. It is difficult for a psychologist to evaluate it from that standpoint and the present review from a psychological viewpoint will naturally be less favorable. It draws its material mainly from biographies, with nothing in the way of actual measurements or ratings by associates.

The work is divided into five parts. The first is introductory, defines leadership in terms of exerting special influence over a number of people and suggests several classifications such as direct-indirect, partizan-scientific, autocratic-democratic. The second deals with heredity draw-

ing considerable material from H. S. Jennings and includng a rather naive discussion of endocrine balance as related to achievement.

In Part III on social stimuli the author is obviously more at home. The influence of geographical and cultural factors and of associates is clarified by numerous pertinent examples from the lives of famous persons. There is however an occasional uncritical presentation of some scientific study that needs interpreting; for instance the fact that unglaciated areas of Indiana produced a smaller proportion of prominent persons than the glaciated regions.

The fourth section is the longest of the book and deals with personality. Here again the approach is anecdotal. Topics include forecasts in childhood, intelligence, *i.e.*, "energy at work," character, *i.e.*, "integraton of personality traits," inhibition, tact, humor and system. Aside from passing mention of Hartshorne and May's work there is practically no reference to psychological sources.

The final section on principles and theories gives a straightforward presentation of introversion-extroversion followed by such topics as focalization of psychic energy, flashes of insight, balance, saturation, *i.e.*, the tendency for one individual to lead too many things and break down. The author makes a plea for democratizing leadership. *i.e.*, passing it around.

The appendices give lists of leaders based on student votes or space in the encyclopedia, a scheme for rating leadership on which the psychologists could improve and a rather useful list of 100 leaders with brief mention of several of the outstanding things each one did. Each chapter is followed by references and a set of exercises for the student which fit in well with the content of the chapter.

The book has little that will interest the psychologist. There is little reference to anything experimental or to any of the techniques for measuring or rating. One who was making a job analysis of leadership as a basis for an experimental program might pick up a few hints from the analysis given of biographical material. Probably Appendix C would be the most useful portion of the work in this connection. The terminology is loose from a psychological standpoint, *e.g.*, "inferiority reflexes," "personal magnetism," "psychic energy," "sudden mental leaps."

The book probably would get some students mildly interested in problems of leadership and might have a slight inspirational flavor. A youth reading biographies for culture or inspiration might find it profitable to read this book in the midst of his program in order to pull together the loose ends of his own thinking and have a little more systematic conception of leadership as revealed in the lives of famous persons.

HAROLD E. BURTT,
Ohio State University.

NEW BOOKS AND PAMPHLETS RECEIVED

Books and pamphlets for review should be mailed to JAMES P. PORTER, *Editor*, JOURNAL OF APPLIED PSYCHOLOGY, Ohio University, Athens, Ohio.

Begabung Im Lichte der Eugenik. DR. JOSEF SOMOGYI. Franz Deuticke, Leipzig und Wien, 1936. 518 pp.

Elementary General Psychology. SAMUEL W. FERNBERGER. F. S. Crofts & Co., 1936. $3.00. 445 pp.

Elgin Papers. Vol. II. Published from the Elgin State Hospital, Elgin, Illinois. April, 1936. 275 pp.

General Psychology: An Outline. DOUGLAS FRYER and EDWIN R. HENRY. Barnes and Noble, New York City, 1936. 75 cents. 229 pp.

Psychology, Elements. KNIGHT DUNLAP. C. V. Mosby Company, St. Louis, 1936. 499 pp.

The Psychology of Feeling and Emotion. CHRISTIAN A. RUCKMICK. McGraw-Hill Book Company, New York, 1936. $4.50. 529 pp.

Psychology and Modern Problems. MORRIS GINSBERG, C. G. SELIGMAN, RAMSAY MUIR, H. CREIGHTON-MILLER, J. C. FLUGEL, EMANUEL MILLER. Longmans-Green & Co., New York, 1936. $2.00. 241 pp.

Readings in Mental Hygiene. ERNEST R. GROVES and PHYLLIS BLANCHARD. Henry Holt & Co., New York, 1936. $2.75. 596 pp.

Reducing the Variability of Supervisor's Judgments: An Experimental Study. SISTER M. XAVIER HIGGINS, R.S.N. Johns Hopkins University Studies in Education, No. 23. Johns Hopkins Press, Baltimore, Md., 1936. 67 pp.

The Return to Religion. HENRY C. LINK. The Macmillan Company, New York, 1936. $1.75. 181 pp.

Social Psychology. ELLIS FREEMAN. Henry Holt & Co., New York, 1936. $2.50. 491 pp.

Tohoku Psychologica Folia. T. CHIBA and OTHERS. Tohoku Imperealis Universitas. Sendai, Japan, 1936. 45 pp.

University Testing Bureau: A Manual of Staff Procedures. E. G. WILLIAMSON, Director, University of Minnesota Press, 1936. $1.00. 50 pp.

Vision et Professions. R. BONNARDEL. Conservatoire National des Arts et Metiers. Paris, IIIe, France, 1936. 25 fr. 164 pp.

Directory of American Psychological Periodicals

(*Continued*)

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research. Founded 1922.

Genetic Psychology Monographs—Worcester, Mass.: Clark University Press. Subscription $7.00 per volume. One volume per year. Edited by Carl Murchison. Bi-monthly. Each number one complete research. Child behavior, animal behavior and comparative psychology. Founded 1925.

Psychological Abstracts—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $6.00. 700 pages ann. Edited by W. S. Hunter and R. R. Willoughby. Monthly. Abstracts of psychological literature. Founded 1927.

Pedagogical Seminary and Journal of Genetic Psychology—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1,000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Child behavior, animal behavior, and comparative psychology. Founded 1891.

Journal of Social Psychology—Worcester, Mass.: Clark University Press. Subscription $7.00. 500-600 pages annually. Edited by John Dewey and Carl Murchison. Quarterly. Political, racial, and differential psychology. Founded 1929.

Personnel Journal—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 500 pages, approximately. Edited by Walter V. Bingham. Bi-Monthly.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research. Founded 1925.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Edward A. Pace. Published without fixed dates, each number a single research. Founded 1926.

Journal of Educational Psychology—Baltimore, Md.: Warwick & York. Subscription $6.00. Edited by Jack W. Dunlap, Harold E. Jones, and Percival M. Symonds. Published monthly except June to August. Founded 1910.

Child Development—Baltimore, Md.: Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Buford Johnson. Quarterly. Founded 1930.

Character and Personality—Durham, N. C.: Duke University Press. Subscription $3.00 per annum. Quarterly. Edited by Robert Saudek (London) and Ernest Seeman (Durham, N. C.). Founded 1932.

The Psychological Exchange—11 East 125th Street, New York, N. Y. Subscription $3.00 per annum. Bi-monthly. Edited by James Hargan. Founded 1932.

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.: Cornell University. Subscription $6.50. 624 pages annually. Edited by K. M. Dallenbach, M. F. Washburn, M. Bentley, E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

Journal of General Psychology—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1,000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Psychological Review—Princeton, N. J.: Psychological Review Company. Subscription $5.00. 540 pages annually. Bi-monthly. General Psychology. Edited by Herbert S. Langfeld. Founded 1894.

Psychological Monograph—Princeton, N. J.: Psychological Review Company. Subscription $6.00 per vol. 500 pp. Ed. by Joseph Peterson. Published without fixed dates, each issue one or more researches. Founded 1895.

Psychological Index—Princeton, N. J.: Psychological Review Company. Subscription $4.00. 300–400 pp. Edited by W. S. Hunter and R. R. Willoughby. An annual bibliography of psychological literature. Founded 1895.

Psychological Bulletin—Princeton, N. J.: Psychological Review Company. Subscription $6.00. 720 pages annually. Psychological literature. Monthly (10 numbers). Edited by John A. McGeoch. Founded 1904.

Archives of Psychology—New York: Columbia University. Subscription $6.00. 500 pp. ann. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $5.00. Foreign, $5.25. Editor, Harry T. Moore. Quarterly. 448 pages annually. Abnormal and Social. Founded 1906.

Psychological Clinic—Philadelphia: Psychological Clinic Press. Subscription $3.00. 288 pages. Ed. by Lightner Witmer. Without fixed dates (Quarterly). Orthogenics, psychology, hygiene. Founded 1907.

Psychoanalytic Review—Washington, D. C.: 3617 10th St., N.W. Subscription $6.00. 500 pages annually. Psychoanalysis. Quarterly. Ed. by W. A. White and S. E. Jelliffe. Founded. 1913.

Journal of Experimental Psychology—Princeton, N. J. Psychological Review Company. 900 pages annually. Experimental. Subscription $7.00. Bi-monthly. Edited by S. W. Fernberger. Founded 1916.

Journal of Applied Psychology—Lancaster, Pa.: Science Press Printing Co. Subscription $6.00. 600 pages annually. Bi-monthly. Edited by James P. Porter, Ohio University, Athens, Ohio. Founded 1917.

Journal of Comparative Psychology—Baltimore: The Williams & Wilkins Co. Subscription $5.00 per volume of 450 pages. Two volumes every year. Edited by Knight Dunlap and Robert M. Yerkes. Founded 1921.

(Continued on Third Cover)